

From: Li, Ruixiang  
Sent: Tuesday, April 16, 2002 4:13 PM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search of Application NO: 09/758,593

Please do a standard search of SEQ ID NOs: 1 and 2 against both the commercial and interference nucleic acid databases.

Thank you very much!

Ruixiang Li  
GAU 1646  
CM1 10E18  
Mail Box 10C01  
306-0282

Point of Contact:  
Toby Port  
Technical Info. Specialist  
CM1 6A04  
703-308-3534

SEQ ID NO. 1 329 aa

Fragment of SEQ ID NO: 2 1158 NT  
576 NT  
253 NT  
569 NT  
330 NT  
SEQ ID NO: 2  
SEQ ID NO: 3  
SEQ ID NO: 4  
SEQ ID NO: 5  
SEQ ID NO: 6

RECEIVED  
APR 16 2002  
STIC

Variant  
Fragment of SEQ ID NO: 2 255 NT  
275 NT  
315 NT  
207 NT  
SEQ ID NO: 7  
8  
9  
10

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 4/17  
Date Completed: 4/18  
Searcher Prep/Review: 10  
Clerical: \_\_\_\_\_  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: 1  
AA Sequences: 1  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST(where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: eg  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

THE  
UNITED STATES  
DEPARTMENT OF  
COMMERCE  
BUREAU OF  
ECONOMIC ANALYSIS  
WASHINGTON, D. C. 20540

THIS PAGE BLANK (USPTO)

**TITLE** Direct Submission  
**JOURNAL** Submitted (23-MAR-2001) Masatsugu Moriyama, Tottori University,  
 Department of Molecular Biology, Faculty of Medicine, Nishimachi  
 86, Yonago, Tottori 683-8503, Japan  
 (E-mail: moriyama@grape.med.tottori-u.ac.jp, Tel: 81-859-34-8040,  
 Fax: 81-859-34-8274)

**FEATURES**  
 source  
 1. .1169  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="TEL"  
 /cell\_type="esophageal carcinoma"  
 /country="Japan:Tottori, Yonago"  
 11. .1012  
 /gene="arpp"  
 11. .1012  
 /gene="arpp"  
 /codon\_start=1  
 /product="ankyrin-repeat protein"  
 /protein\_id="BAB60958.1"  
 /db\_xref="GI:14495178"  
 /translation="MDQTMEDSEAVORATALIEORLAQFEENEKLRGDTROKLPMDLL  
 VLDEKHGSAOLOKVGOERVKTSLDLREIIDVGIQNLIELKRRKOKRDA  
 LAAHEPPEPEITGVDETEFLKAVEGKMVIERFLADGGSADTCDQFRITLHR  
 ASLEHMEILEKLDNGATVDFODRLDCTAMHWACRGHLEVVKLKLSHGADTNVRDK  
 LLSTPLHVAVTGQTEIVEHFLSIGLEINARDREGDTALHDVRLNRYKIILKLLHG  
 ADMWTKNLAKTDTDLVLQWQADTRHLEHPEPEHNGLEPNDSGRETPQVPVPAQ"  
 BASE COUNT 279 a 321 c 383 g 186 t  
 ORIGIN

Query Match 99.3%; Score 1149.4; DB 9; Length 1169;  
 Best Local Similarity 99.9%; Pred. No. 2e-211;  
 Matches 1150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 aggcacgacccatggaggaactcagagcggtgcagagggccacagcgctcatcgagcag 67  
 DB 11 ATGACGGCACCATGAGGAGCTCCGAGCGGTGCAGAGGGCCACAGCGCTCATCGAGCAG 70  
 QY 68 cggctggcacagggagggagaatgagaactccgagagagacacacgcagaaagtgcgc 127  
 DB 71 CGGCTGGCACAGGAGGAGGAATGAGAACTCCGAGGAGACACACGCCAAGACTGCC 130  
 QY 128 atggaactgtgtgtcgtgaggaatgagaagcacacagggggtcagatgcagcctgcag 187  
 DB 131 ATGACTTGTGCTGTGAGGATGAGAAGCACACCGGGGCTCAGAGTGCAGCCTGCAG 190  
 QY 188 aagtgaaaggccaaagacgctgcgaagacgtccctggacctgcggcgaggatcatc 247  
 DB 191 AAGGTGAAGGCCAAGACGCGCTGCGAAGACGTCCTCGACCTCGCGCGGGAGATCATC 250  
 QY 248 gatgtggcgggatccagaacctcatcgagctgcggaagaaacgcagacagaagaacgg 307  
 DB 251 GATGTGGCGGGATCCAGAACTCATCGAGCTCGGAAGAAACCAAGCAAGAAAGCGG 310  
 QY 308 gacgctctggcgctcgcagatgacgccccccagagccccaggagatcatcgtggcctgtg 367  
 DB 311 GACGCTCTGGCGGCTCGCATGAGCGCGCCCCACAGCCCGAGGAGACACTGGCCCTGTG 370  
 QY 368 gatgagagacctctcgaagagctgcggtggaggggaaatgaaggctattgagaagttc 427  
 DB 371 GATGAGGAGACCTTCCGTAAAGCTGCGGTGGAGGGGAAAATGAAGGTCATTGAGAAGTTC 430  
 QY 428 ctggctgacgggggttcagccgacacgtgcgaccagttccgtcgacagcagctcaccga 487  
 DB 431 CTGGCTGACGGGGGTTCAGCCGACACGTGCACAGTCCGTCGAGACGACTGCACCGCA 490  
 QY 488 gcttccctggaagggccacatggaatcctggagagcttctagataatggggccacttg 547  
 DB 491 GCTTCCCTGGAAGGCCACATGGAATCCTGGAGAAGCTTCTAGATAATGGGGCCACTGTG 550  
 QY 548 gacttcaggaatcggtggaactgcacagcctgcatctggcgctccgcgggggccaactta 607

DB 551 GACTTCCAGGATCGGCTGGACTGCACAGCCCATGATTGGGCTCCGCGGGGGCCACTTA 610  
 QY 608 gaggtgtgaaactctgcaaaagccatggagcagacacacacacacacacacacacacacac 667  
 DB 611 GAGTGTGTAACCTTCGCAAGCCATGAGGACACACACACACACACACACACACACACAC 670  
 QY 668 agcaccctcgtgcagctggcagtcgacagggcgaggtggagattgtggagcactttcta 727  
 DB 671 AGCAACCCCGCTGCACCTGGCAGTCCGGACAGGGCAGGTGGAGATTGTGGAGCACTTCTA 730  
 QY 728 tccctgggctggaaatacaatccagagagagagagagagagagagagagagagagagagag 787  
 DB 731 TCCCTGGGCTTCGAAATCAATGCCAGAGACAGGAGGAGGAGGAGTACTGCCCTCATGACGCT 790  
 QY 788 gtgaggtcacaaccgctcaaaatcatcaaaactcgtctccctgcagatgggctgcacatgatg 847  
 DB 791 GTGAGGCTCAACCCGCTACAATTCATCAACTGCTCTCCATGAGGGGCTGCATGATG 850  
 QY 848 accaagaacctggcaggaagacccctcagggacctggtgcagctctggcagcgtcgataacc 907  
 DB 851 ACCAAGAACCTGGCAGGAGAAAGACCCCGACGACCTGGTGCAGCTTCGCGAGCTGATACC 910  
 QY 908 cggcacgctcggagatcctcctgagccgggggctgagcataaacgagcggctgagggcctaat 967  
 DB 911 CGGCACGCTGGAGCATCTCTGAGCCGGGGCTGAGCAATAACGGGCTGGAGGGGCTAAT 970  
 QY 968 gatagtggcgagagacccctcagcctcgtccagccagtgaaatgcgtgccccagccag 1027  
 DB 971 GATAGTGGCGAGAGACCCCTCAGCCTGTGCCACCCAGTGAATGCTTGCCTCCAGCCGAG 1030  
 QY 1028 ccagctaccagccgccccctctctgtgtgcagcgggggtcctaaagaatggctcccgagct 1087  
 DB 1031 CCAGCTACCCAGGCCCCCTCTGTGTGCAGCGGAGGGTCTCTAAGAATGGCTCCCGAGCT 1090  
 QY 1088 aactgagggggccagcctttttctgcatgatccagggagacacatacccaaacaccacaa 1147  
 DB 1091 AACTGAGGGGCCAGCCCTTTTCTGTCATGATCCAGAGGACATACCAAACTACCACAA 1150  
 QY 1148 taaaaaagctg 1158  
 DB 1151 TAAAAAAGCTG 1161

**RESULT 2**  
**LOCUS** HSA304805 1159 bp mRNA PRI 21-DEC-2000  
**DEFINITION** Homo sapiens mRNA for skeletal muscle ankyrin protein 2 (ANKRD2 gene).  
**ACCESSION** AJ304805  
**VERSION** AJ304805.1 GI:11967782  
**KEYWORDS** ANKRD2 gene; skeletal muscle ankyrin protein 2.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE** 1 (bases 1 to 1159)  
**AUTHORS** Pallavicini, A., Kojic, S., Bean, C., Vainzof, M., Salamon, M.,  
 Ievolella, C., Bortoletto, G., Pacchioni, B., Trevisan, S.,  
 Faulkner, G., Lanfranchi, G. and Valie, G.  
**TITLE** Characterization of human skeletal muscle Ankrd2  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 1159)  
**AUTHORS** Pallavicini, A.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (18-DEC-2000) Pallavicini A., Biologia, Università di  
 Padova, via Ugo Bassi 58/b, 35100, ITALY  
**FEATURES**  
 source  
 1. .1159  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 11. .1144  
 /gene="ANKRD2"  
 11. .1012

**THIS PAGE BLANK (USPTO)**



QY	788	gtgaggctcaacgcgtacaataatcatcaactgctgctcctgcatggggctgacatgatg	847
Dd	790		
Dd	791	GTGAGGCTCAACCGCTACAAAATCATCAAAGTGGCTGCTCCTGCATGGGGCTGCATGATG	850
QY	848	accaagaacctggcaggaagaccgccagcgagacctggtgagctcttggcaggctgatacc	907
Dd	851		
Dd	851	ACCAAAGAACCTGGCAGGAAGAACCCGACGACACTGGTGCACTCTGGCAGCTCTGGCAGGCTGATACC	910
QY	908	cggcacgccccctggagcatccttgagccgggggctgagcataaacgggctggaggggcctaatt	967
Dd	911		
Dd	911	CGGCACGCCCTGGAGCATCTTCGACCGGGGCTGAGCATACGGGCTGGAGGGGCGCTTAAT	970
QY	968	gatatggggcagagagaccccctcagcctgtgccagcccgagtgaatcgctgccccagccccag	1027
Dd	971	GATAGTGGGGCAGAGACCCCTCAGCCTGTGCCACCCAGTGAATCGCTGSCCAGCCCCAG	1030
QY	1028	ccagctaccagccccctctctgtgtgagccggggggtcctaaagaatggctcccgaggct	1087
Dd	1031		
Dd	1031	-CCAG-----CTCTCTGTGTGGACCGCGGAGGTTCTTAAGATGGTCTCCGGAGCT	1079
QY	1088	aactgagggccagcccttttttcgcatgattcagagagacacataccaaaactaccacaa	1147
Dd	1080		
Dd	1080	AACGTAGGGGCCAGCCCTTTTTCTGCATGATCCAGGAGCACATACCAAACTACCCAA	1139
QY	1148	tataaaaa 1154	
Dd	1140	ATAAAA 1146	

RESULT 3

```
MMU011118      1100 bp      mRNA          ROD      19-MAR-2001
LOCUS          Mus musculus mRNA for a skeletal muscle and cardiac protein.
DEFINITION     AJ011118
ACCESSION      AJ011118
VERSION        AJ011118.1 GI:5420271
KEYWORDS       skeletal muscle and cardiac protein.
SOURCE         house mouse.
ORGANISM       Mus musculus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 1100)
AUTHORS        Kempy,T.J., Sadusky,T.J., Saltisi,F., Carey,N.N., Moss,J., Yang,S.Y.,
                Sassoon,D.A., Goldspink,G. and Coulton,G.R.
TITLE          Identification of Ankrd2, a novel skeletal muscle gene coding for a
                stretch-responsive ankyrin-repeat protein
JOURNAL        Genomics 66 (3), 229-241 (2000)
MEDLINE        20334618
REFERENCE      2 (bases 1 to 1100).
AUTHORS        Ievolella,C., Formentin,E. and Lanfranchi,G.
TITLE          Characterization of a member of a new family proteins with ankyrin
                repeats
JOURNAL        Unpublished
REFERENCE      3 (bases 1 to 1100)
AUTHORS        Ievolella,C.
TITLE          Direct Submission
COMMENT        Submitted (30-SEP-1998) Ievolella C., CRIBI Biotechnology Centre,
                Universita di Padova, via G. Colombo 3 Padova, 35121, ITALIA
FEATURES       Related sequence F18281.
                Location/Qualifiers
                1..1100
                 /organism="Mus musculus"
                 /db_xref="taxon:10090"
                 /tissue_type="diaphragm muscle"
                 /clone_lib="Uni-ZAP TM XR, Stratagene"
                 <1..10
                  /evidence=experimental
                 11..997
                  /codon_start=1
                  /product="skeletal muscle and cardiac protein"
                  /protein_id="CA546646.1"
                  /db_xref="GI:5420272"
                  /sb_xref="SPTRMBL:Q9WV06"
                  /translation="MECPFAVORATELIEORLAOEETENRRSAAPKLSMDMLVLVEE"
```

5'UTR

CDS

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2002, 08:10:37 ; Search time 26.67 Seconds  
(without alignments)  
1804.409 Million cell updates/sec

Title: US-09-758-593A-1  
Perfect score: 1693  
Sequence: 1 MEDSAVQRATALIQRALQ.....NGLEGNDSGRETPQVPQAQ 329

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL17.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1693	100.0	333	4 Q9GZV1	Q9GZV1 homo sapien
2	1484.5	87.7	328	11 Q9WV06	Q9WV06 mus musculus
3	677	40.0	319	6 Q9TU71	Q9TU71 oryctolagus
4	669.5	39.5	319	11 Q9Z1F0	Q9Z1F0 rattus norv
5	667	39.4	319	11 Q501A4	Q501A4 mus musculus
6	666	39.3	319	11 Q9CR42	Q9CR42 mus musculus
7	658	38.9	319	4 Q153Z7	Q153Z7 homo sapien
8	505	29.8	102	4 Q9NQC9	Q9NQC9 homo sapien
9	276	16.3	1214	5 Q9VUX2	Q9VUX2 drosophila
10	249.5	14.7	1327	4 Q95271	Q95271 homo sapien
11	248.5	14.7	307	6 Q9N043	Q9N043 macaca fasc
12	247.5	14.6	543	10 Q9LOP7	Q9LOP7 arabidopsis
13	247.5	14.6	1166	4 Q9H2K2	Q9H2K2 homo sapien
14	247.5	14.6	1265	4 Q9HAS4	Q9HAS4 homo sapien
15	244	14.4	1050	4 Q9H014	Q9H014 homo sapien
16	241.5	14.3	592	12 Q9J4Z4	Q9J4Z4 fowlpox vir
17	241.5	14.3	1136	6 Q9N180	Q9N180 bos taurus
18	240.5	14.2	548	10 Q9LVG7	Q9LVG7 arabidopsis
19	239.5	14.1	1549	5 Q9V4B1	Q9V4B1 drosophila

20 238.5 14.1 1549 5 Q24241 Q24241 drosophila  
21 238.5 14.1 1715 11 Q9ERD4 Q9ERD4 rattus norv  
22 238.5 14.1 1763 11 Q9EOG6 Q9EOG6 rattus norv  
23 237.5 14.0 247 10 Q9Z079 Q9Z079 arabidopsis  
24 235 13.9 1031 4 Q9UF42 Q9UF42 homo sapien  
25 234.5 13.9 1719 4 Q13768 Q13768 homo sapien  
26 234.5 13.9 1856 4 Q99407 Q99407 homo sapien  
27 234 13.8 1159 5 Q9NCP8 Q9NCP8 drosophila  
28 234 13.8 1430 11 Q9JJF7 Q9JJF7 mus musculu  
29 234 13.8 2443 5 Q9VSA2 Q9VSA2 drosophila  
30 233.5 13.8 1848 11 Q61302 Q61302 mus musculu  
31 231.5 13.7 843 11 P97582 P97582 rattus norv  
32 231 13.6 271 4 Q9NPL9 Q9NPL9 homo sapien  
33 231 13.6 1181 5 Q9XZ37 Q9XZ37 drosophila  
34 231 13.6 1181 5 Q9VBP3 Q9VBP3 drosophila  
35 230 13.6 411 11 Q9D4D8 Q9D4D8 mus musculu  
36 230 13.6 1943 11 Q61307 Q61307 mus musculu  
37 229 13.5 4377 4 Q12955 Q12955 homo sapien  
38 228.5 13.5 318 13 Q91974 Q91974 gallus gall  
39 228.5 13.5 2119 5 Q9VAU5 Q9VAU5 drosophila  
40 228 13.5 637 5 Q9VUW9 Q9VUW9 drosophila  
41 228 13.5 762 5 Q9VVR3 Q9VVR3 drosophila  
42 228 13.5 1088 4 Q13484 Q13484 homo sapien  
43 226.5 13.4 272 10 Q9AUW0 Q9AUW0 oryza sativ  
44 226 13.3 627 10 Q9SYK5 Q9SYK5 arabidopsis  
45 226 13.3 669 4 Q9BSK4 Q9BSK4 homo sapien

#### ALIGNMENTS

RESULT 1  
Q9GZV1  
ID Q9GZV1 PRELIMINARY; PRT; 333 AA.  
AC Q9GZV1  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE SKELETAL MUSCLE ANKYRIN PROTEIN 2 (SKELETAL MUSCLE ANKYRIN REPEAT).  
GN ANKRD2  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fallavacini A., Kojic S., Bean C., Vainzof M., Salamon M.,  
RA Ivoletta C., Bortolotto G., Pacchioni B., Trevisan S., Faulkner G.,  
RA Lanfranchi G., Valle G.;  
RL "Characterization of human skeletal muscle Ankrd2";  
DR EMBL; AJ304805; CAC19412.1;  
DR EMBL; AJ304804; CAC19411.1;  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 4.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 5  
DR PROSITE; PS50088; ANK\_REPEAT; 4.  
DR PROSITE; PS50297; ANK\_REPEAT; 1.  
SQ SEQUENCE 333 AA; 37151 MW; 679736F0491467A8 CRC64;

Query Match 100.0%; Score 1693; DB 4; Length 333;  
Best Local Similarity 100.0%; Pred. No. 3e-109;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEDSAVQRATALIQRALQAEENEKLRGDTKQLPMDLLVLEDEKHGAQSAALQKVG 60

Db 5 MEDSAVQRATALIQRALQAEENEKLRGDTKQLPMDLLVLEDEKHGAQSAALQKVG 64

QY 61 QERVRKTSLDLRREITDVGGIQNLIELRRKKRKRDLAALASHEPPPEETITGPVDEET 120

Db 65 QERVRKTSLDLRREITDVGGIQNLIELRRKKRKRDLAALASHEPPPEETITGPVDEET 124

**THIS PAGE BLANK (USPTO)**

QY 121 FLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFOD 180  
 Db 125 FLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFOD 184  
 QY 181 RLDCTAMHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGL 240  
 Db 185 RLDCTAMHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGL 244  
 QY 241 EINARDREGDTALHDAVRLNRYKIIKLLHGHADMNTKLNAGKTPDVLVQLWQADTRHAL 300  
 Db 245 EINARDREGDTALHDAVRLNRYKIIKLLHGHADMNTKLNAGKTPDVLVQLWQADTRHAL 304  
 QY 301 EHPEPGAENHGLEPNDSGRETPOQVPAQ 329  
 Db 305 EHPEPGAENHGLEPNDSGRETPOQVPAQ 333

RESULT 2  
 QY06 PRELIMINARY; PRT; 328 AA.  
 QY06:  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE SKELETAL MUSCLE AND CARDIAC PROTEIN (ANKYRIN REPEAT DOMAIN 2).  
 GN ANKRD2  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-DIAPHRAGM MUSCLE;  
 RA Ievoliella C., Formentin E., Lanfranchi G.;  
 RT "Characterization of a member of a new family proteins with ankyrin repeats.";  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VAIOUS STRAINS; TISSUE-SPLEEN, AND SKELETAL MUSCLE;  
 RX MEDLINE=20334618; PubMed=10873377;  
 RA Kemp T.J., Sadusky T.J., Saltis F., Carey N., Moss J., Yang S.Y.,  
 RA Sassoon D.A., Goldspink G., Coulton G.R.;  
 RT "Identification of Ankr2, a novel skeletal muscle gene coding for a stretch-responsive ankyrin-repeat protein.";  
 RL Genomics 66:229-241(2000). ID=6157200  
 EMBL; AJ011118; CAB46646.1; -;  
 EMBL; AJ245346; CAB99432.1; -;  
 EMBL; AJ245514; CAB99431.1; -;  
 HSP; P25963; IINK.  
 DR MGD; MGI:1861447; Ankr2.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; ank; 4.  
 DR SMART; SM00248; ANK; 4.  
 DR PROSITE; PS50088; ANK REPEAT; 4.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 SQ SEQUENCE 328 AA; 36707 MW; DB90D955EE9D175E CRC64;

Query Match 87.7%; Score 1484.5; DB 11; Length 328;  
 Best Local Similarity 88.8%; Pred. No. 7.1e-95;  
 Matches 292; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 1 MEDSEAVORATLIEORLAQEEENELKRGDTROKLPMDLLVLEDEKHGQAASALQKVGK 60  
 Db 1 MEGPEAVORATELIEORLAQEEETELRRSAPGLSMDMLVLEEKRLGVQSPALQKVGK 60  
 QY 61 QERVKTSLDLREIIVDVGIONLIELRKKRQKRDALAAASHEPPEEITGPVDEET 120  
 Db 61 QERVKTSLDLREIIVDVGIONLIELRKKRQKRDALAAQAEPPPEEITGPVNEET 120  
 QY 121 FLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFOD 180

Db 121 FLKAAVEGKMKVIDYKLLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFOD 180  
 QY 181 RLDCTAMHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGL 240  
 Db 181 RLDCTAMHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGL 240  
 QY 241 EINARDREGDTALHDAVRLNRYKIIKLLHGHADMNTKLNAGKTPDVLVQLWQADTRHAL 300  
 Db 241 DINARDREGDTALHDAVRLNRYKIIKLLHGHADMNTKLNAGKTPDVLVQLWQADTRHAL 300  
 QY 301 EHPEPGAENHGLEPNDSGRETPOQVPAQ 329  
 Db 301 EHPEPGAENHGLEPNDSGRETPOQVPAQ 328

RESULT 3  
 QY071 PRELIMINARY; PRT; 319 AA.  
 AC QY071:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE CARP  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20011195; PubMed=10542334;  
 RA Aihara Y., Kurabayashi M., Arai M., Kedes L., Nagai R.;  
 RT "Molecular cloning of rabbit CARP cDNA and its regulated expression in adriamycin-cardiomyopathy.";  
 RL Biochim. Biophys. Acta 1447:318-324(1999).  
 DR EMBL; AF131883; AAF3817.1; -;  
 DR HSP; Q00421; IAWC.  
 DR InterPro; IPR002110; ANK.  
 DR Pfam; PF00023; ank; 4.  
 DR SMART; SM00248; ANK; 4.  
 DR PROSITE; PS50088; ANK\_REPEAT; 4.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
 SQ SEQUENCE 319 AA; 36245 MW; C7EF87A02936FED3 CRC64;

Query Match 40.0%; Score 677; DB 6; Length 319;  
 Best Local Similarity 46.4%; Pred. No. 3.2e-39;  
 Matches 136; Conservative 54; Mismatches 69; Indels 34; Gaps 3;

QY 29 GDTROKLPMDLLVLEDEKHGQAASALQKVGKE-----RV 64  
 Db 18 GETGEFLP-----DDFRDQGYEAAVTSEKEDLKLTPAHVSVLAEOQWEREKLEAEL 70  
 QY 65 RKTSLDLRREIIVDVGIONLIELRKKRQKRDALAAASHEPPEEITGPVDEETFLKA 124  
 Db 71 KKKKLEQSKLENLEDELEIIQLKREKRYKTKVPVAKL---PEPIITEPVDVPEFLKA 127  
 QY 125 AVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFODRLDC 184  
 Db 128 ALENKLVAVYEKFLSDQNNPDVDEYKRTALHRACLEGLAIVEKLMGAQAIEFRDMLRS 187  
 QY 185 TAMHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGLEINA 244  
 Db 188 TAIHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGLEINA 247  
 QY 245 ROREGDTALHDAVRLNRYKIIKLLHGHADMNTKLNAGKTPDVLVQLWQADTR 297  
 Db 248 KOREGDTPLHDAVRLNRYKIIKLLHGHADMNTKLNAGKTPDVLVQLWQADTR 297

RESULT 4  
 QY21F0 PRELIMINARY; PRT; 319 AA.  
 ID QY21F0

**THIS PAGE BLANK (uspto)**



**THIS PAGE BLANK (uspto)**



```

Qy 74 gcacaggaggaggaatgagaaactcgcagagagagacacacacccaggaagctgcccacgac 133
Db 5 GCACAGGAGGAGGAGGAATGAGAAATCCGAGGAGACACACCCAGAGCTGCCATGGAC 64
Qy 134 ttgctgggtcgtggagatgagaagaccacacgggctcagatgcagccctcagaagtg 193
Db 65 TTGCTGGTCTGGAGGATGAGAAGCACACCGGGGCTCAGATGTCAGCCCTGCAGAAAGTG 124
Qy 194 aaggcccaagagcgtgcgcaagacgtccctggacctgcgcgggagatcatcgatgtg 253
Db 125 AAGGCCAAGAGCGGTGCGCAAGACGCTCCCTGGACCTGCGCGGAGATCATCGATGTG 184
Qy 254 ggcgggattccgaacctcatcgatcgtcggaagaacacgagcagaagaagcggacgct 313
Db 185 GCGGGATCCGAACCTCATCGATGCGGAGAAACGCAAGCAAGCAAGCGGACGCT 244
Qy 314 ctggccgctcgtcagcagcccccagagcccgagagatcactgcccctgtgatgag 373
Db 245 CTGGCCGCTCGATGAGCCGCCCCAGAGCCGAGGAGATCACTGGCCCTGTGGATGAG 304
374 gag 376
Db 305 GAG 307

:RESULT 6
HSA304804
LOCUS HSA304804 14300 bp DNA PRI 21-DEC-2000
DEFINITION Homo sapiens ANKRD2 gene for skeletal muscle ankyrin repeat, exons
1-9.
ACCESSION AJ304804
VERSION AJ304804.1 GI:11967780
KEYWORDS ANKRD2 gene; skeletal muscle ankyrin repeat.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 14300)
AUTHORS Pallavicini A., Kojic S., Bean C., Vainzof M., Salamon M.,
Ievolella C., Bortoletto G., Pacchioni B., Trevisan S.,
Faulkner G., Lanfranchi G. and Valle G.
TITLE Characterization of human skeletal muscle Ankrd2
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 14300)
AUTHORS Pallavicini A.
TITLE Direct Submission
JOURNAL Submitted (18-DEC-2000) Pallavicini A., Biologia, Universita di
Padova, via Ugo Bassi 58/b, 35100, ITALY
URES Location/Qualifiers
source 1. 14300
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="10q23.31-23.32"
1538. 12679
/gene="ANKRD2"
1538. 1542
/gene="ANKRD2"
1567. 1563
/gene="ANKRD2"
/number=1
join(1567. 1563,6596. 6697,7036. 7194,7294. 7401,
9651. 9749,10112. 10210,11111. 11209,11400. 11498,
12373. 12679)
/gene="ANKRD2"
join(1577. 1663,6596. 6697,7036. 7194,7294. 7401,
9651. 9749,10112. 10210,11111. 11209,11400. 11498,
12373. 12522)
/gene="ANKRD2"
/codon_start=1
/product="skeletal muscle ankyrin repeat"
/protein_id="CAC19411.1"
/db_xref="gi:11967781"

```

```

/translation="MDGTMEDSEAVQRATALIEQRLAQEEENEKLRDGTQKLPMDLL
VLDEKHGQSAALQKVGOERVRKTSLDLREIIDVGGTQNLIELRKKRQKRRDA
LAASHEPPPEEITGVDEETFLKAAVEGKMVKIEFLADGGSDTCDOFRRTALHR
ASLEGHMEILEKLLDNGATVDFODLCTAMHWACRGHLEVKLLQSHGADTNVRDK
LLSTPLHAVRTGQVEIVEHFLSLGLEINARDREGDTRALHDVRLNRYKIKLLLLHG
ADMTKNLAGKTPTDLVLQWQADTRHALEHPEPGAENHLEGGPNDSGRETPOPPVPAQ"
1664. 6595
/gene="ANKRD2"
/number=1
6596. 6697
/gene="ANKRD2"
/number=2
6698. 7035
/gene="ANKRD2"
/number=2
7036. 7194
/gene="ANKRD2"
/number=3
7195. 7293
/gene="ANKRD2"
/number=3
7294. 7401
/gene="ANKRD2"
/number=4
7402. 9650
/gene="ANKRD2"
/number=4
9651. 9749
/gene="ANKRD2"
/number=5
9750. 10111
/gene="ANKRD2"
/number=5
10112. 10210
/gene="ANKRD2"
/number=6
10211. 11110
/gene="ANKRD2"
/number=6
11111. 11209
/gene="ANKRD2"
/number=7
11210. 11399
/gene="ANKRD2"
/number=7
11400. 11498
/gene="ANKRD2"
/number=8
11499. 12372
/gene="ANKRD2"
/number=8
12373. 12679
/gene="ANKRD2"
/number=9
12659. 12664
/gene="ANKRD2"
BASE COUNT 3383 a 3611 c 3952 g 3353 t 1 others
ORIGIN

```

```

Query Match 26.0%; Score 301.4; DB 9; Length 14300;
Best Local Similarity 99.7%; Pred. No. 1.3e-48;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 856 cctggcaggaagaccccgagcctgtgctgagctgtgagcagctgataccggcagc 915
Db 12369 CCAGCGAGGAAGACCCGAGCGACCTGTGTCAGCTCTGGCAGGTGATACCCGCGC 12428
Qy 916 cctggagcatcctgagccgggggctgagcataaacgggctgaggggccaatgatagc 975
Db 12429 CCTGGAGCATCTGTAGCCGGGGCTGAGCATAACCGGGCTGAGCCTAATGATAGTGG 12488
Qy 976 gcgagagaccctcagcctgtgccagcccagtcgaatgcgtgcccccagcccagcagctac 1035

```

THIS PAGE BLANK (USPTO)

```
Db 12489 GCAGAGACCCCTCAGCTGTCCAGCCAGTGAATGCTGCCAGCCAGCCAGCTAC 12548
QY 1036 ccagccctctgtgtgacgaggggtcctaagaatgctcccgagctaaactaggg 1095
Db 12549 CCAGCCCTCTGTGTGACGCGAGGGTCTTAAGATGCTCCCGAGGTAAGTGGG 12608
QY 1096 gccagccttttctcatgatccagagcacataccacaaataccacataaaaaag 1155
Db 12609 GCCAGCCTTTTCTCATGATCCAGGACATACACAACTACCACATAAAAAAG 12668
QY 1156 ctg 1158
Db 12669 CTG 12671
```

```
RESULT 7
AL355315 163214 bp DNA HTG 07-AUG-2001
LOCUS Homo sapiens chromosome 10 clone RP11-548K23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
ACCESSION AL355315
VERSION AL355315.12 G&:15141465
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163214)
Chapman, J.
Direct Submission
Submitted (06-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 9, 2001 this sequence version replaced gi:15041819.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA548K23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 162481 bases at least Q40
Consensus quality: 162660 bases at least Q30
Consensus quality: 162811 bases at least Q20
Insert size: 163014; sum-of-contigs
Insert size: 169230; agarose-fp
Quality coverage: 10.60x in Q20 bases; sum-of-contigs Quality
coverage: 10.65x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 42225: contig of 42225 bp in length
* 42226 42325: gap of 100 bp
* 42326 139511: contig of 97186 bp in length
* 139512 139611: gap of 100 bp
* 139612 163214: contig of 23603 bp in length.
Location/Qualifiers
1. 163214
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-548K23"
```

FEATURES  
source

```
misc_feature /clone_lib="RPC1-11.2"
1. 42225
/note="assembly fragment: 05104
fragment_chain:1"
misc_feature 42326..139511
/note="assembly fragment: 02594
fragment_chain:1"
misc_feature 139612..163214
/note="assembly fragment: 00919
fragment_chain:1"
BASE COUNT 41214 a 38587 c 39191 g 44022 t 200 others
ORIGIN
Query Match 26.0% Score 301.4; DB 2; Length 163214;
Best Local Similarity 99.7%; Pred. No. 7.2e-49;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 856' cctggaggaagaccacccgacggacgtggtgagctctgagcagctgatacccgagcgc 915
Db 12347 CCAGGCAGAAAGACCCGACGACCTGGTGCAGCTCTGGCAGGCTGATACCCGCGCAGC 12406
QY 916 cctggagcactcctgagccgggggctgagcataacggtgaggggctgaggggctaatgatgtg 975
Db 12407 CCTGGAGCATCTCAGCGGGGGCTGAGCATAGCGGGCTGAGGGGGCTAATGATAGTGG 12466
QY 976 gcgagagaccctcagcctgtgccagccagctgaatgctgccccagccagccagctac 1035
Db 12467 GCGAGAGACCCCTGAGCTGTGCCAGCCAGTGAATGCTGCCAGCCAGCCAGCCAGTAC 12526
QY 1036 ccagccctctctgtgacgagggggtgagcaggaatgctcccgagctaaactaggg 1095
Db 12527 CCAGCCCTCTCTGTGACGCGAGGGTCTTAAGATGCTCCCGAGGCTAAGTGG 12586
QY 1096 gccagccttttctcatgatccagagcacataccacaaataccacataaaaaag 1155
Db 12587 GCCAGCCTTTTCTCATGATCCAGGACATACACAACTACCACATAAAAAAG 12646
QY 1156 ctg 1158
Db 12647 CTG 12649
RESULT 8
AL359388 218959 bp DNA HTG 02-AUG-2001
LOCUS Homo sapiens chromosome 10 clone RP11-445I23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 15 unordered pieces.
ACCESSION AL359388
VERSION AL359388.21 GI:15041841
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 218959)
Clark, S.
Direct Submission
Submitted (01-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 30, 2001 this sequence version replaced gi:15022209.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA445I23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 162481 bases at least Q40
Consensus quality: 162660 bases at least Q30
Consensus quality: 162811 bases at least Q20
Insert size: 163014; sum-of-contigs
Insert size: 169230; agarose-fp
Quality coverage: 10.60x in Q20 bases; sum-of-contigs Quality
coverage: 10.65x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 42225: contig of 42225 bp in length
* 42226 42325: gap of 100 bp
* 42326 139511: contig of 97186 bp in length
* 139512 139611: gap of 100 bp
* 139612 163214: contig of 23603 bp in length.
Location/Qualifiers
1. 163214
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-548K23"
```

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 08:09:16 ; Search time 1721.2 Seconds  
(without alignments)  
11099.078 Million cell updates/sec

Title: US-09-758-593A-2

Perfect score: 1158

Sequence: 1 cagctcgaggacgcaccca.....ctaccacataaaaagctg 1158

coring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Maximum number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pi.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_om.\*

20: em\_or.\*

21: em\_ov.\*

22: em\_pat.\*

23: em\_ph.\*

24: em\_pl.\*

25: em\_ro.\*

26: em\_sts.\*

27: em\_sy.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htgo\_hum.\*

31: em\_htgo\_inv.\*

32: em\_htgo\_rod.\*

33: em\_htg\_hum.\*

34: em\_htg\_inv.\*

35: em\_htg\_rod.\*

36: em\_htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1149.4	99.3	1169	9	AB058599 Homo sapi
2	1109.8	95.8	1159	9	AJ304805 Homo sapi
3	761.1	65.7	1100	10	MM0011118
4	759.4	65.6	1100	10	MM0245514
5	302.2	26.1	307	9	AJ245975 Homo sapi
6	301.4	26.0	14300	9	AJ304804 Homo sapi
7	301.4	26.0	163214	2	AL355315 Homo sapi
8	301.4	26.0	218959	2	AL359388 Homo sapi
9	264.8	22.9	1758	6	AX140502 Sequence
10	261.6	22.6	1026	10	AF041847 Mus muscu
11	256.8	22.2	1749	10	U50736 Rattus norv
12	248.4	21.5	1901	9	HSRNOACINP
13	248.4	21.5	1901	11	G28603
14	243.6	21.0	1940	4	AF131883
15	131.8	11.4	9876	10	MM0249346
16	113.4	9.8	386	6	AX071347 Sequence
17	103.8	9.0	363	6	AX071332 Sequence
18	91.6	7.9	222	10	AF041849
19	88.2	7.6	6192	9	HUMANK
20	87.2	7.5	5080	9	AB023174 Homo sapi
21	86.6	7.5	7252	9	HSANKY
22	86.6	7.5	197748	2	AC026283
23	85.8	7.4	6278	9	HSBRANK2
24	85.8	7.4	12518	9	HSANKB440
25	85	7.3	67923	9	AL391244 Human DNA
26	84.6	7.3	2754	9	AB028932 Homo sapi
27	84.6	7.3	3370	9	BC011608 Homo sapi
28	84.6	7.3	17173	9	AB058779 Homo sapi
29	82.6	7.1	7084	10	MUSANK3B
30	82.6	7.1	7140	10	AF239045
31	81.8	7.1	3452	4	AF222766 Bos tauru
32	80.6	7.0	2529	10	RU065916
33	79.4	6.9	6207	10	AF069525 Rattus no
34	79.4	6.9	7867	10	AF102552 Rattus no
35	78.6	6.8	3520	10	MMANKIAC
36	78.6	6.8	6238	10	MUSANKIA
37	78.6	6.8	8145	10	MMANKIAA
38	78.6	6.8	40143	10	AF185591
39	78.4	6.8	2562	1	CVNFCCAB
40	77.6	6.7	1742	9	AB046089
41	77.4	6.7	1934	5	AF296130 Gallus ga
42	77.2	6.7	1135	9	AF105427 Homo sapi
43	77.2	6.7	1288	6	AR088277 Sequence
44	77.2	6.7	1345	6	AX025001 Sequence
45	77.2	6.7	1345	9	AF094760 Homo sapi

ALIGNMENTS

RESULT 1

AB058599

LOCUS

DEFINITION

AB058599

AB058599

AB058599.1

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (sites)

Moriyama,M., Tsukamoto,Y., Fujiwara,M., Kondo,G., Nakada,C.,

Baba,T., Ishiguro,N., Miyazaki,A., Nakamura,K., Hori,N., Sato,K.,

Shomori,K., Takeuchi,K., Sato,H., Mori,S. and Ito,H.

Identification of a novel human ankyrin-repeated protein homologous

to CARP

Biochem. Biophys. Res. Commun. (2001) In press

REFERENCE

2 (bases 1 to 1169)

Moriyama,M.

Thu Apr 18 10:26:05 2002

TITLE Direct Submission  
 JOURNAL Submitted (23-MAR-2001) Masatsugu Moriyama, Tottori University,  
 Department of Molecular Biology, Faculty of Medicine, Nishimachi  
 86, Yonago, Tottori 683-8503, Japan  
 (E-mail: moriyama@grape.med.tottori-u.ac.jp, Tel: 81-859-34-8040,  
 Fax: 81-859-34-8274)  
 FEATURES Location/Qualifiers  
 source 1..1169  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /cell\_line="YEL"  
 /cell\_type="esophageal carcinoma"  
 /country="Japan:Tottori, Yonago"  
 11..1012  
 /gene="arpp"  
 11..1012  
 /gene="arpp"  
 /codon\_start=1  
 /product="ankyrin-repeat protein"  
 /protein\_id="BAB60958.1"  
 /db\_xref="GI:14495178"  
 /translation="MDGTMEDSAVQRATALIEQRJAEEENKLRGTRQKLPMDLL  
 VLEKHHGASALOKVGOERVKTSRLRREIIDVGGIQLIELRKRKKRDA  
 LAASHEPPPEETGTGVDRETFKAAVEGKMKVIEKFLADGSDTCDFRRTALHR  
 ASLEGHMELKRLDNGATVDPRDCTAHMWACRGHLEVVKLLQSHGADTNVRDK  
 LLSTPLHVAVRTGQVEIHFELSLGLEINARDREGDTALDVLNRYKIILKLLHG  
 ADMMTKNLAGTKPTDLVOLWQADTRHALRHPFGAENHGLEPNDSRETPQVPAQ"  
 BASE COUNT 279 a 321 c 383 g 186 t  
 ORIGIN  
 Query Match 99.3%; Score 1149.4; DB 9; Length 1169;  
 Best Local Similarity 99.9%; Pred. NO. 2e-211;  
 Matches 1150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 YQ 8 agggacgcacatgagggactcgagggcggtgagggcgccacagcgctcatcgagcg 67  
 DB 11 ATGACGCGACCATGAGGAGACTCCGAGCGGTGCAGAGGCGCCACAGCGCTCATCGAGCAG 70  
 YQ 68 cggctggcacagggaggaagaatgagaaactccgagggagacacagccagagctgccc 127  
 DB 71 CGGCTGGCACAGGAGGAGGAGATGAGAACTCCGAGGAGACACAGCCAGAGGTGCC 130  
 YQ 128 atgaactgctggtgctggaggtgagaaagcaccacaggggctcagagtgagcgccctgag 187  
 DB 131 ATGGACATTGCTGCTGGAGGATGAGAACACACAGCGGCTCAGAGTGACGCTGCAG 190  
 YQ 188 aaggtgaaggcccaagagcgctgccaagacgtccctgaacctgcggcgaggagatcacc 247  
 DB 191 AAGGTGAAGGGCCAAAGAGCGCGTCCGCAAGACGTCCTTGACCTGCGCGGGAGATCATC 250  
 YQ 248 gatgtggcgagatccagaaactcatgctgcggaagaacgaagcagaagaagcgcg 307  
 DB 251 GATGTGGCGGGATCCAGAACTCATGAGTGCAGGAAGAACCAAGCAGAGAGAGCGG 310  
 YQ 308 gacgctctggcgctcgcagtagccgccccagagcccgaggagatcactggcctgtg 367  
 DB 311 GACGCTCTGGCGGCTCGCATGAGCGCGCCCGCAGAGCCCGAGGAGATCATCTGGCCCTGTG 370  
 YQ 368 gatgaagacacctctgaagctcggtggaggggaaatgaagggtcattgagaagttc 427  
 DB 371 GATGAGGAGACCTTCTGAAAGCTGGGTGGAGGGGAAATGAAGGTGATGAGAAGTTC 430  
 YQ 428 ctggctgacggggggtcagccacacgtgacacaggtccctcgagacagactcaccga 487  
 DB 431 CTGGCTGACGGGGGTGACGCGACACGTGCGACCAAGTTCCTCGTGGACAGACTGCACCGA 490  
 YQ 488 gcttcctggaagggacacatgaaatcctggagaagctcttagaataatggggcactgtg 547  
 DB 491 GCTTCCCTGGAAGGCCACATGGAATCTCTGGAGAGCTTCTAGATAATGGGGCCACTGTG 550  
 YQ 548 gacttcagagatcggtgagctgcacagccatgcatgtggcctgcgcggggggcactta 607

Db 551 GACTTCAGGATCGGCTGACTGCACAGCCATGCATTGGCCTGCGCGGGGGCCACTTA 610  
 YQ 608 gagtggtgaaacttctgcaaaagccatggagagacacacaaatgtgagggataagctgctg 667  
 Db 611 GAGTGTGAACTTCTGCAAGCCATGGAGCAGACACCAATGTGAGGGATAGCTGTG 670  
 YQ 668 agcaccctgcgtgacgtggcagtcgagccagagcaggtgagattgtgagcactteta 727  
 Db 671 AGCACCCTGCTGACGTGCGAGTCCGACAGGCGAGGTGGAGATTGTGGAGCATTCTCA 730  
 YQ 728 tccctggcctgaaatcaatgccagagacaggaaggagatactgccccctgacgct 787  
 Db 731 TCCCTGGGCTTGGAAATCATGCGAGAGACAGGAGGGGATACTGCCCTGCTGACGCT 790  
 YQ 788 gtgaggtcaaccgctaaataatcatcaactgctgctcctgcatggggctgacatgag 847  
 Db 791 GTGAGGCTCAACCGCTAGAAATCATCAAACTGCTCTCTGATGGGGCTGACATGATG 850  
 YQ 848 accaagaacctggcaggaagacaccccgacgacctggtgacgtctgagcaggtgatacc 907  
 Db 851 ACCAAGAACCCTGGCAGGAAAGACCCCGACGACCTGCTGGCAGCTCTGGCAGGCTGATAC 910  
 YQ 908 cggcagcctctgagcatcctgagccgggggctgagcagcagcagcagcagcagcagc 967  
 Db 911 CGGCAGCCCTGGAGCATCCTGAGCGGGGCTGAGCATAACGGGCTGGAGGGCCTAAT 970  
 YQ 968 gatgtggcgagagacccctcagcctgctgcccagccagtgaaatgctgccccagccag 1027  
 Db 971 GATAGTGGGCGAGAGACCCCTCAGCCTGTGCCAGCCAGTGAATGCTGCTGCCAGCCAG 1030  
 YQ 1028 caagctaccacgcccctctctgtgtgcagccggagggtcctaagaatgctccccgagct 1087  
 Db 1031 CCAGCTACCCAGCCCTCTCTGTGTGCAGCGGAGGGCTCTAAGAAATGGCTCCCGAGCT 1090  
 YQ 1088 aactgagggccccagccttttctcatgacagcagcagcagcagcagcagcagcagcagc 1147  
 Db 1091 AACTGAGGGCCAGCCCTTTTCTGCTGATCCAGGAGCAGATACACAAACTACCAAA 1150  
 YQ 1148 taaaaagctg 1158  
 Db 1151 TAAAAAGCTG 1161  
 RESULT 2  
 HSA304805 1159 bp mRNA PRI 21-DEC-2000  
 LOCUS Homo sapiens mRNA for skeletal muscle ankyrin protein 2 (ANKRD2  
 DEFINITION gene)  
 ACCESSION AJ304805  
 VERSION AJ304805.1 GI:11967782  
 KEYWORDS ANKRD2 gene; skeletal muscle ankyrin protein 2.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1159)  
 AUTHORS Pallavicini A., Kojic S., Bean C., Vainzof M., Salamon M.,  
 Ievoli G., Bortoletto G., Pacchioni B., Trevisan S.,  
 Faulkner G., Lanfranchi G. and Valle G.  
 TITLE Characterization of human skeletal muscle Ankrd2  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1159)  
 AUTHORS Pallavicini A.  
 JOURNAL Direct Submission  
 TITLE Submitted (18-DEC-2000) Pallavicini A., Biologia, Università di  
 Padova, via Ugo Bassi 58/b, 35100, ITALY  
 FEATURES Location/Qualifiers  
 source 1..1159  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 11..1144  
 /gene="ANKRD2"  
 11..1012  
 gene  
 CDS

[illegible]

Qy	788	gtaggctcaacgcgtacaaaaatcatcaaaactgctgctctctgcatggggctgacatgatg	847
Db	791	GTGAGGCTCAACCCGCTACAAAATCATCAAACTGCTGCTCTGTCATGGGGCTGACATGATG	850
Qy	848	accaagaacctggcagaaagaccacagacctggtgcagctctggcagcctgatacc	907
Db	851	ACCAAGAACCTGGCAGAAAGACCCCGACGACCTGGTGCAGCTCTGCGAGGCTGATACC	910
Qy	908	cggcacgacctggagcatcctgagccgggggctgagcataaacgggctggaggggacctaat	967
Db	911	CGGCACCCCTGGAGCATCTGACCGGGGCTGAGCATTAACGGGCTGGAGGGCCCTAAT	970
Qy	968	gatagtggcgagagacccctcagctgtgtccagcccaagtgaatcgctgccccagccccag	1027
Db	971	GATAGTGGCGAGAGACCCCTCAGCCTGTGCAGCCAGCCAGTGAATGCGTGCSCCAGCCAG	1030
Qy	1028	ccagctaccagccccctctctgtgtgagccggagggtcctaaagaatgtctcccggaact	1087
Db	1031	CCAG-----CTCTCTGTGTGCAGCCGAGGGTCTTAAAGTAATGGCTCCCGGAGCT	1079
Qy	1088	aactgagggccagccctttttctgcatgcatccaggagcacatcaccaaaactaccacaa	1147
Db	1080	Aactgagggccagccctttttctgcatgcatccaggagcacatcaccaaaactaccacaa	1139
Qy	1148	taaaaa	1154
Db	1140	ATAAAA	1146
RESULT 3			
MMU011118			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
COMMENT			
FEATURES			
source			
5'UTR			
CDS			

Query Match	65.7%	Score 761	DB 10	Length 1100
Best Local Similarity	82.1%	Pred. No. 9.8e-137		
Matches 888	Conservative 0	Mismatches 190	Indels 3	Gaps 17
QY	10	ggacggccacatggaggactccgagggcggtgcagagggccacacgctcatcgagcagcg	69	
DB	1	GGAGGGTACCATGGAGGGGCCGAGGCGGTGACAGAGAGCCACAGAGCTATCAGCAGCG	60	
QY	70	gctggccacagaggaggagaatgagaaactccgagagacacacgcgcagagctgccccat	129	
DB	61	GCTTGCCAGGAGGAGAGACTGAGAACTTCGAAGATCCGCTCCTGGGAGGCTGTCCAT	120	
QY	130	gcaattgctggctgtaggagatgagaagcaccacgggctcagagtcgacccctcgagaa	189	
DB	121	GGACATGCTAGTGTAGGAGAGAGAGCGCTCGGGGTTTCAGAGTCTGCTTTACAAA	180	
QY	190	ggtgaaggcccaagagcgtgctgcgaagctccctggaacctgcggcgaggagatcatcga	249	
DB	181	GTTAAGGGCCAGAGCGCGTGCGAAGACATCCCTGGACCTGGCGGTGAGATCATTGA	240	
QY	250	tgtggcgggatccagaacctcatcgctgcgaagaacgcgaagcagaagaagcggga	309	
DB	241	TGTGGGTGGGATCCAGAACTCATTTGAAGTGAAGGAAAACGCAAGCAGAAAAGCGAGA	300	
QY	310	cgctctggccctcgcatgagccgccccagagcccgaggagatcaactggccctgtgga	369	
DB	301	TGCCCTGGCTGCAGCCAGAGGCTCTCCAGAGCAGAGGAGATCACTGGCCCTGTGAA	360	
QY	370	tgaggagaccttccataaagctcggtgaggggaaaaatgaagtcattgagaagtccct	429	
DB	361	TGAGGAGACATTCCTGAAAGCTGCATGTGGAGGGGAAAATGAAAGTCATTGCAAGTACCT	420	
QY	430	ggctgacggggctcagcgcaacgtgcacacagttccgtcgacagcagcactgcaccgagc	489	
DB	421	GGCGGAGCGAGGTTACGGGACACCTGTGATGNGTCCGTCGGACAGCACATGCATCGGGC	480	
QY	490	tccctggaagggccacatggaaatcctggagaagctctagataatggggccactgtgga	549	
DB	481	CTCCTGGAGGACACATGGAGATCTGGAGAACTCTGGAGAAATGGGGGCCACCGTGA	540	
QY	550	cttcaggatcggtgactgcacagccatgcattggcctgcgcggggggccacttaga	609	
DB	541	CTTCAGGATCGGGTGGACTGACAGCCCATGCACTGGGCGTGCCTGGCGTGGGGGCCACCTGA	600	
QY	610	ggtggtgaacctctgcaagccatggagcagacacccaatgtgagggataagctgctgag	669	
DB	601	GGTGTGTGAGACTTCCTGCNAAGCCGGGGGGCCGACACAACTGAGAGACAAGCTACTGAG	660	
QY	670	caccgcgctgcactggcagtcgccggacagggcgaggtgagattgtgaggaactttatc	729	
DB	661	CACCTCCCTGCATGTGGCGCTCCGTACTTGGACACGTGGAGATTGTGGAACAWTTTCTTTC	720	
QY	730	cttgggctggaatcaatgcacagacaggaaggagatactgcccctgcatgacgctgt	789	
DB	721	CTTGGGCTTGGATATCAATGTCGAAGACAGAAAGGGACAGTGCCTTACATCATGCCGT	780	



```

AKNLGKTPDLDVLQWQADTRHALEHPEPESEQNGLRPGSGRETPOPIPAQ"
repeat_region 455..850
/rpt_family="Ankyrin"
/rpt_type=TANDEM
3'UTR 998..1100
/polyA_signal 1081..1086
/polyA_site 1100
BASE COUNT 284 a 294 c 341 g 181 t
ORIGIN
Query Match 65.6%; Score 759.4; DB 10; Length 1100;
Best Local Similarity 82.1%; Pred. No. 2e-136;
Matches 88; Conservative 0; Mismatches 191; Indels 3; Gaps 1;
Qy 10 ggaagggcaccatgagggactccgagggcgtgagggcggcagcagcgtcatcgagcagc 69
Db 1 GGAGGGTACCATGAGGGGCGCGAGGGGTGCAGAGAGCCACAGAGCTCATCGAGCAGCG 60
Qy 70 gctggcagagagagagagagagagagagagagagagagagagagagagagagagagat 129
Db 61 GCTTGGCCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Qy 130 ggaacttgctgctgagggatgagaagcaccacgggctcagagtcagagtcagccctgcagaa 189
Db 121 GCACATGCTAGTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 190 ggtgaagggcacaagagcgtgctgcaagacgtccctggacctggcggggagagatcaga 249
Db 181 GGTGAAGGCCCAAGAGCGCGTGCAGACACATCCCTGGACCTGCAGCTGAGATCATTTGA 240
Qy 250 tgtggcgggatccagaacctcatcgagtcggaagaaagcagaagcagaagcagcagc 309
Db 241 TGTGGTGGGATCCAGAACCTCATTTGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
Qy 310 cgccttgccgctcgtgagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 369
Db 301 TGCCCTTGCTGCAGCCAGGAGGCTCTCTCCAGAGCCAGAGGAGATCAGCGGCCCTGTGA 360
Qy 370 tgaagagacctctctgaagctgctgagggcgggggagagagagagagagagagagagag 429
Db 361 TGAGAGAGACATTCCTGAAAGTGCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Qy 430 ggcctgacgggggtcagcgcagcgtgcgaccagtcctcgtcgagcagcagcagcagcagc 489
Db 421 GCGGACGGAGGTTGAGCGGACACCTGTGATGAGTTCGTCGAGCAGCAGCAGCAGCAGCAG 480
Qy 490 ttccctgaaaggccacatggaatcctggagaagcttcttagataatggggccactgtgga 549
Db 481 CTCCCTGGAGGAGACATGGAGATGAGTGGAGAACTTCTGGAGATGGGGCCACCGTGA 540
Qy 550 cttccagagatcggtgagctgcacagccatgcattggcctgcccgcggggccacttaga 609
Db 541 CTTCCAGGATCGGCTGGAGCTGCACAGCCATGCACCTGGGCTGCCCTGGGGGCCACCTGA 600
Qy 610 ggtggtgaaactctgaaagccatggagcagacacccaatgtgagggataagctcgaag 669
Db 601 GGTGTGAGACTCTCTGAAAGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 660
Qy 670 caccgcgtgcagctggcagtcgcagacagggcaggtggagagtgtggagcactttctatc 729
Db 661 CACTCCCTGCATGTGGCGTCCGATCTGGACACCTGGAGAGATTGTGGAACATTTCTTTC 720
Qy 730 cctgggctggaatcaatgcagagagagagagagagagagagagagagagagagagagag 789
Db 721 CCGTGGCTTGGATATCAATGCCAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Qy 790 gaggctcaacgcgtcaaaaatcaatcaactgctgctcgtcgtgagggcgtgacatgagc 849
Db 781 GAGACTCAACCGCTCAAAAATCATCAAACTGCTGCTTGTGATGGGGCGGAGATGATGCG 840

```

```

Qy 850 caaqaacctggcagaaagaccacacgagacctggtgagctctgagcaggtgatacccg 909
Db 841 TAAGAACCTGGCGGAAAGACACCCACGAGACCTGTCTCAGCTCTGGCAGAGACACCG 900
Qy 910 gcagccctgagcatcctgagcgggggctgagcataacgggctggaggcctaatga 969
Db 901 GCATGCCCTGGAGCATCCGAGCCAGAATCAGAGCAGAACGAGCTGGAGAGGCC---TGG 957
Qy 970 tagtgggagagagaccctcagcgtgtccagccagcagtgaaatgctgcccagccagcc 1029
Db 958 GAGTGGCGTGAGACACCTCAGCCTTATACCGCCAGTAAATACCTACCCGAGAGCTGGGC 1017
Qy 1030 agctaccagcgcctctctgtgctgagcgggaggtcctaaagaatggctcccgagctaa 1089
Db 1018 ACCCAACCCCTACTACTATGAGCTGCTGCAGCCAGCTCTAGCAGGAACCCCTAGATGCAA 1077
Qy 1090 c 1090
Db 1078 c 1078

RESULT 5
HSA249975 307 bp mRNA PRI 25-JUL-2000
LOCUS Homo sapiens partial mRNA for ankyrin repeat domain 2 (stretch responsive muscle).
ACCESSION AJ249975
VERSION AJ249975.1 GI:9501289
KEYWORDS Ankr2 gene; ankyrin repeat domain 2; stretch responsive muscle.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 307)
AUTHORS Kemp, T.J., Sadusky, T.J., Saltisi, F., Carey, N., Moss, J., Yang, S.Y., Sassoon, D.A., Goldspink, G. and Coulton, G.R.
TITLE Identification of ankr2, a novel skeletal muscle gene coding for a stretch-responsive ankyrin-repeat protein
JOURNAL Genomics 66 (3), 229-241 (2000)
MEDLINE 20334618
AUTHORS Kemp, T.J.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1999) Kemp T.J., Molecular Pathology, Imperial College School Of Medicine, 3AF Building Level 2, Exhibition Road, South Kensington, London, UNITED KINGDOM
COMMENT Related sequence: AJ245514.
FEATURES
Location/Qualifiers
1..307
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="skeletal muscle"
1..307
/gene="ANKRD2"
<1..>307
/gene="ANKRD2"
/note="stretch responsive muscle"
/codon_start=2
/product="ankyrin repeat domain 2"
/protein_id="CAB99416.1"
/db_xref="GI:9501290"
/translation="SAOEENEKLGRDXRKLPMDDLVLLEDEKHHGAQSAALQKVKGO
ERVRKTSLDLRREIIVDVGIGLNLIELRKRKRKRDALAAASHPEEITGPVDEE
"
BASE COUNT 76 a 82 c 110 g 38 t 1 others
ORIGIN
Query Match 26.1%; Score 302; DB 9; Length 307;
Best Local Similarity 99.7%; Pred. No. 2.5e-48;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Thu Apr 18 10:26:05 2002

QY	74	gcacaggaggagagaatgagaactccgagagagacacacgcccagaagctgcccatggac	133
Db	5	GCACAGGAGGAGGAGATGAGAACTCCGAGGAGACNCACGCCAGAACTGCCCATGGAC	64
QY	134	ttgctggtgctgagatgagaagcaccacggggtcagatgcagccctgcagaaggtg	193
Db	65	TTGCTGGTCTGGAGGATGAGAAGCACACGGGGCTCAGAGTGCAGCCCTGCAGAAGGTG	124
QY	194	aaggcccaagcgcgtgcgaagacgtccctggacctggcggcgggagatcatcgatgtg	253
Db	125	AAGGGCAAGAGCGCGTGGCAGACGTCCTGTGGACCTCGCGGGGAGATCATCGATGTG	184
QY	254	ggcggatccagaacctcatcgctgcggaagaacgcaagcagaagaagcggagcgt	313
Db	185	GGCGGGATCCAGAACTCATCGAGTGGCGGAGAACGCAACAGAAAGCGGACGCT	244
QY	314	ctggccqctcgcagtagcgcgcacagccccagagcccgaggagatcactggccctgtgatgag	373
Db	245	CTGGCGCCTCGCATGAGCGGCCCGCAGAGCCCGAGGAGATCACTGGCCCTGTGGATGAG	304
QY	374	gag 376	
Db	305	GAG 307	
RESULT 6			
LOCUS	HSA304804	14300 bp	DNA
DEFINITION	Homo sapiens ANKRD2 gene for skeletal muscle ankyrin repeat, exons 1-9.		
ACCESSION	AJ304804		
VERSION	AJ304804.1		
KEYWORDS	ANKRD2 gene; skeletal muscle ankyrin repeat.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 14300)		
TITLE	Pallavicini, A., Kojic, S., Bean, C., Vainzof, M., Salamon, M., Ievolella, C., Bortoletto, G., Paccioni, B., Trevisan, S., Faulkner, G., Lanfranchi, G. and Valle, G.		
JOURNAL	Characterization of human skeletal muscle Ankr2		
FEATURES	Unpublished		
source	2 (bases 1 to 14300)		
REFERENCE	Pallavicini, A.		
AUTHORS	Direct Submission		
TITLE	Submitted (18-DEC-2000) Pallavicini A., Biologia, Universita di Padova, via Ugo Bassi 58/b, 35100, ITALY		
JOURNAL	Location/Qualifiers		
1..14300	/organism="Homo sapiens"		
/db_xref="taxon:9606"	/chromosome="10"		
/map="10q23.31-23.32"	1538..12679		
/gene="ANKRD2"	1538..1542		
/gene="ANKRD2"	1567..1663		
/gene="ANKRD2"	/number=1		
Join(1567..1663,6596..6697,7036..7194,7294..7401,9651..9749,10112..10210,11111..11209,11400..11498,12373..12679)	/gene="ANKRD2"		
Join(1577..1663,6596..6697,7036..7194,7294..7401,9651..9749,10112..10210,11111..11209,11400..11498,12373..12522)	/gene="ANKRD2"		
/codon_start=1	/product="skeletal muscle ankyrin repeat"		
/protein_id="CAC19411.1"	/db_xref="GI:11967781"		
/translation="MDGTMEDSEAVORATALIEORLAQBEENEKLRGDTROKLPMDLLVLEDEKHGAQSAALQKVGOERVRKTS			
LDRREIIVDVGIGIQLNIELKKRCKKRD			
LAASHEPPPEPEITGVDEETFLKAAGEKMKVIEKFLADGGSADTCQDFRTALHRSASLSEHMEILEKLLDNGATVDFQDRLDCTAMHWACRGHLEVVKLQSHGADTNVADKLLSTPLHAVVTRTGQVEIVHFSLSLGLEINARDREGDTALHDAVRLNRYKIKIQLLLHGADMTKKNLAKGTPDILVOLWQADTRHALEHPEPCAEHNGLEGPNDSGRETPQVPAQ"			
intron	1664..6595		
	/gene="ANKRD2"		
	/number=1		
exon	6596..6697		
	/gene="ANKRD2"		
	/number=2		
intron	6698..7035		
	/gene="ANKRD2"		
	/number=2		
exon	7036..7194		
	/gene="ANKRD2"		
	/number=3		
intron	7195..7293		
	/gene="ANKRD2"		
	/number=3		
exon	7294..7401		
	/gene="ANKRD2"		
	/number=4		
intron	7402..9650		
	/gene="ANKRD2"		
	/number=4		
exon	9651..9749		
	/gene="ANKRD2"		
	/number=5		
intron	9750..10111		
	/gene="ANKRD2"		
	/number=5		
exon	10112..10210		
	/gene="ANKRD2"		
	/number=6		
intron	10211..11110		
	/gene="ANKRD2"		
	/number=6		
exon	11111..11209		
	/gene="ANKRD2"		
	/number=7		
intron	11210..11399		
	/gene="ANKRD2"		
	/number=7		
exon	11400..11498		
	/gene="ANKRD2"		
	/number=8		
intron	11499..12372		
	/gene="ANKRD2"		
	/number=8		
exon	12373..12679		
	/gene="ANKRD2"		
	/number=9		
polya_signal	12659..12664		
	/gene="ANKRD2"		
BASE COUNT	3383 a	3611 c	3952 g
ORIGIN	1 others		
Query Match 26.0%; Score 301.4; DB 9; Length 14300;			
Best Local Similarity 99.7%; Pred. No. 1.3e-48;			
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	856	cctggcaggaagacccccagacacctgtgtgacgtctctggcaggtgatacccgacgc	915
Db	12369	CCAGGCAGGAAGAACCCGACGACCTGTGTGACCTGTGGAGGCTGATACCCGCGACGC	12428
QY	916	cctggagacatcctgagccggggctgagcacaacgggctggagggcctaataatgattgg	975
Db	12429	CCTGGAGCATCCTGAGCCGGGGCTGAGCATAACGGGCTGGAGGGCCCTAATGATAGTG	12488
QY	976	gcgagagacccctcagcgtgtgcagcccagtcgaatgctgtgcccagcccagccagctac	1035

```
|||||
Db 12489 GCGAGAGACCCCTCAGCTGTGCCAGCCAGTGAATGCGTGTGCCAGCCAGCCAGCTAC 12548
QY 1036 ccagccctctgtgtcacccgaggttcttaagaatgctccgagtaactgagg 1095
Db 12549 CCAGCCCTCTCTGTGTGTCAGCCGAGGGTCTTAAGATGGCTCCCGAGCTAACTGAGG 12608
QY 1096 gccagccttttttgcgtatcatcagagcacataccacaaactaccacataaaaaag 1155
Db 12609 GCCAGCCCTTTTCTGTCATGATCCAGGAGCACATACCACAACTACCCACATAAAAAAG 12668
QY 1156 ctg 1158
Db 12669 CTG 12671
|||
```

```
RESULT 7
AL355315 163214 bp DNA HTG 07-AUG-2001
LOCUS Homo sapiens chromosome 10 clone RP11-548K23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 unordered pieces.
ACCESSION AL355315
VERSION AL355315.12 GI:15141465
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163214)
Chapman, J.
Direct Submission
Submitted (06-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 9, 2001 this sequence version replaced gi:15041819.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba548K23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 162483 bases at least Q40
Consensus quality: 162860 bases at least Q30
Consensus quality: 162811 bases at least Q20
Insert size: 163014; sum-of-contigs
Quality coverage: 10,60x in Q20 bases; sum-of-contigs Quality
coverage: 10,65x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 42225: contig of 42225 bp in length
* 42226 42325: gap of 100 bp
* 42326 139511: contig of 97186 bp in length
* 139512 139611: gap of 100 bp
* 139612 163214: contig of 23603 bp in length.
Location/Qualifiers
1. 163214
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-548K23"
```

```
FEATURES
source
```

```
/clone_lib="RPC1-11.2"
1. 42225
/note="assembly fragment:05104"
fragment_chain:1"
42326..139511
/note="assembly fragment:02594"
fragment_chain:1"
139612..163214
/note="assembly fragment:00919"
fragment_chain:1"
BASE COUNT 41214 a 38587 c 39191 g 44022 t 200 others
ORIGIN
```

```
Query Match 26.0%; Score 301.4; DB 2; Length 163214;
Best Local Similarity 99.7%; Pred. No. 7.2e-49;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 856 cctggcagaagaccccgacgacctgtgcagctgtgcagctgtgcagctgtgcagctgtgcagcgc 915
|||
Db 12347 CCAGCAGGAAAGACCCCGACGACCTGGTGCAGCTCTGGCAGGCTGATACCCCGCAGCGC 12406
QY 916 cctggagatcctgagccgggggctgagcataaacgggctggaggggcctaataatgattgg 975
|||
Db 12407 CTTGAGGATCTTGAGCGGGGGCTGAGCATAACGGGCTGGAGGGCCCTAATGATAGTGG 12466
QY 976 gcgagagacccctcagcctgtgcagccagctgaatgcgtgccccagccagccagctac 1035
|||
Db 12467 GCGAGAGACCCCTCAGCCCTGTGCCAGCCAGCTGATGCTGCCAGCCAGCCAGCTAC 12526
QY 1036 ccagccctctgtgtgcagccggggtctctaagaatggtctcccgagctaaatgagg 1095
|||
Db 12527 CCAGCCCTCTCTGTGTGTCAGCGGAGGGTCTTAAGATGGTCTCCGGAGCTAACTGAGG 12586
QY 1096 gccagcctttttctgcatcattccagagcacataccacaaactaccacataaaaaag 1155
|||
Db 12587 GCCAGCCCTTTTCTGTCATGATCCAGGAGCACATACCACAAATACCACATAAAAAAG 12646
QY 1156 ctg 1158
Db 12647 CTG 12649
|||
```

```
RESULT 8
AL359388 218959 bp DNA HTG 02-AUG-2001
LOCUS Homo sapiens chromosome 10 clone RP11-445I23, *** SEQUENCING IN
DEFINITION PROGRESS ***, 15 unordered pieces.
ACCESSION AL359388
VERSION AL359388.21 GI:15041841
KEYWORDS HTG; HTGS_PHASE1; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 218959)
Clark, S.
Direct Submission
Submitted (01-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jul 30, 2001 this sequence version replaced gi:15022209.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba445I23
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
```

SOURCE	house mouse.
ORGANISM	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
REFERENCE	1 (bases 1 to 1758)
AUTHORS	Woll, E., Werner, S., Halle, J.P., Regenbogen, J. and Goppelt, A.
TITLE	Use of polypeptides or their encoding nucleic acids for the diagnosis or treatment of skin diseases and their use in identifying pharmacologically active substances
JOURNAL	Patent: EP 1114862-A 124 11-JUL-2001.
FEATURES	Location/Qualifiers
	1..1758



[illegible]

Qy 458 gaccagttccgtcgacagcactgcacagagcttccctggaggccacatggaatcctg 517  
 Db 697 GATGAGTATAAAGCGACAGCTCTTCATAGAGCATCTTGAAGGACATTTGGCAATTGTG 756  
 Qy 518 gagaagctcttagataaattggggccactgtggaactccaggatcggtgactgcacagcc 577  
 Db 757 GAGAAGTTAATGGAAGCTGGAGCCAGATCGAAATTCCTGTATGCTTGAATCCACAGCC 816  
 Qy 578 atgcattgggctgcgcggggccacttagagtggtgaaactctctgaaagccatgga 637  
 Db 817 ATCCACTGGCAGCCGCTGGAGGAACCTGGATGTTTTAAATTTGCTGATTAAGA 876  
 Qy 638 gcagacaccaatgtgaggataaagctgtgagcaccgcgcgcagctgcagtcggcgaca 697  
 Db 877 GCAAAATATAGCCCGGAGATAAGTTGCTACGACAGCGCTGCATGTGCGGTGAGGACT 936  
 Qy 698 gggcaggtggagattgtgagcaacttctatccctggcctggaaatacaatgcacagac 757  
 Db 937 GGCCCACTATGAGTGGCGGAGCATCTTATCGCCTGTGAGGCAGACCTCAACGCCAAAGAC 996  
 Qy 758 agggagggggactgctccctcatgacgtgtgagcctcaaccgctacacaaatcatcaa 817  
 Db 997 AGAGAGAGAGATACCCGCTGGATGATGGGTGACACTGAACCGTATAGATGATCGCA 1056  
 Qy 818 ctgctgctctctgactggggctgacatgatgacaaagaacctggcaggaaagaccccgag 877  
 Db 1057 CTCTCTGATTATATGCGCGGATCTCAACATCAAGAACTGTGCTGGGAAGACGCGGATG 1116  
 Qy 878 gacctgggtgagctctgagcagctgatacc 907  
 Db 1117 GATCTGGTGTACTACTGGCAGAAATGGAACC 1146

RESULT 13  
 G28603 1901 bp DNA STS 11-JUL-1996  
 LOCUS human STS SHGC-35401, sequence tagged site.  
 DEFINITION G28603  
 ACCESSION G28603  
 VERSION G28603.1 GI:1408418  
 KEYWORDS STS; STS sequence; primer; sequence tagged site.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1901)  
 AUTHORS Myers, R.M.  
 JOURNAL Unpublished (1996)  
 COMMENT  
 Contact: Richard M. Myers  
 Stanford Human Genome Center (SHGC)  
 Stanford University School of Medicine  
 Department of Genetics, M-344, Stanford, CA 94305, USA  
 Tel: 4157259687  
 Fax: 4157259689  
 Email: myers@shgc.stanford.edu  
 Primer A: GGCATTTTGAAGGCATGG  
 Primer B: CCAGATGGATCATCAAGG  
 STS size: 222  
 PCR Profile:  
 Initial incubation: 94 degrees C for 90 seconds  
 Denaturation: 94 degrees C for 15 seconds  
 Annealing: 62 degrees C for 23 seconds  
 Polymerization: 72 degrees C for 30 seconds  
 PCR Cycles: 30  
 Thermal Cycler: Perkin Elmer 9600  
 Protocol:  
 Template: 25 ng  
 Primer: each 1 uM  
 dNTPs: each 200 uM  
 Taq Polymerase: 0.05 units/uL  
 Total Vol: 10 uL

Buffer: MgCl2: 2.5 mM  
 KCl: 50 mM  
 Tris-HCl: 20 mM  
 pH: 8.3  
 Prepared with primer pairs provided by Sandoz, derived from X83703  
 -- Washington University/Merck EST sequence.  
 FEATURES  
 source 1..1901  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="10"  
 STS 1246..1467  
 primer\_bind 1246..1263  
 primer\_bind complement(1447..1467)  
 BASE COUNT 592 a 378 c 460 g 471 t  
 ORIGIN

Query Match 21.5%; Score 248.4; DB 11; Length 1901;  
 Best Local Similarity 64.7%; Pred. No. 3.3e-38;  
 Matches 369; Conservative 0; Mismatches 201; Indels 0; Gaps 0;  
 Qy 338 ccagagccggagagactgacctggtgagagagacacctctctctgaaagctcggtg 397  
 Db 577 CCAGAACCTGAAATCATACGGAACCTGTGGATGTGCCTACGTTCTGGAAGCTGCTCTG 636  
 Qy 398 gaggggaaaaatgaagtcattgagaagttcctggtgacgggggggtgcagccagacgtgc 457  
 Db 637 GAGAATAAATGCCAGTAGTAGAAAAATTTCTGTGACAGCAAGAANAATCCAGATGTTTGT 696  
 Qy 458 gaccagttccgtcgagacagcactgcacagcgttccctggaggccacatggaatcctg 517  
 Db 697 GATGAGTATAAAGCGACAGCTCTTCATAGAGCATCTTGGAAAGGACATTTGGCAATTGTG 756  
 Qy 518 gagaagctcttagataaattggggcactgtgagctccaggatcggtgactgcacagcc 577  
 Db 757 GAGAAGTTAATGGAAGCTGGAGCCAGATCGAATTCCTGTGATATGCTTGAATCCACAGCC 816  
 Qy 578 atgcattgggctgcgcggggccacttagagggtggtgaaactctctgaaagccatgga 637  
 Db 817 ATCCACTGGGAGGCGGTGGAGGAACCTGGATGTTTTAAATTTGTTGCTGAATAAGA 876  
 Qy 638 qcagacaccaatgtgaggataaagctgtgagcaccgcgcagctgcagtcgagtcagcgaca 697  
 Db 877 GCATAAATTAGCCCGGAGATAAGTTGCTCAGCAGACGCTGCATGTGCGGTGAGGACT 936  
 Qy 698 gggcaggtggagatgtgagcactttctatccctgggctggaatcaatccagagac 757  
 Db 937 GGCCCACTATGAGTGGCGGAGCATCTTATCGCTGTGAGGCAGACCTCAACGCCAAAGAC 996  
 Qy 758 agggaaaggggatactgcccctgacgtgtgaggtgcacccgctcaaccgtacataatcaaaa 817  
 Db 997 AGAGAAGGAGATACCCCGTTGCATGATGCGGTGAGACTGAACCGTATTAAGATGCCA 1056  
 Qy 818 ctgctgctctctgaggggctgacatgatgacaaagaacctggcaggaagaccccgagc 877  
 Db 1057 CTCTCTGATTATGATGGCGGATCTCAACATCAAGAACTGTGCTGGGAAGACGCGGATG 1116  
 Qy 878 gacctgggtgagctctgagcagctgatacc 907  
 Db 1117 GATCTGGTGTACTACTGGCAGAAATGGAACC 1146  
 RESULT 14  
 AF131883  
 LOCUS AF131883 1940 bp mRNA MAM 30-NOV-1999  
 DEFINITION Oryctolagus cuniculus CARP mRNA, complete cds.  
 ACCESSION AF131883  
 VERSION AF131883.1 GI:6478316  
 KEYWORDS

Db	1026	CTCCTGATTATGTACGGTCCGACCTCACCATCAGACAGCGCGGAGACCCCAATG	1085
Qy	878	gacctggtgcagctctggcaggctatacc	907
Db	1086	GATCTGGTGTAAACTGGCAAAATGGAACC	1115
RESULT 15			
MMU249346	9876 bp	DNA	ROD
LOCUS	Mus musculus Ankrd2 gene for ankryrin repeat domain 2 (stretch responsive muscle), exons 1-9.		25-JUL-2000
DEFINITION			
ACCESSION	AJ249346		
VERSION	AJ249346.1	GI:9501361	
KEYWORDS	Ankrd2 gene; ankryrin repeat domain 2; stretch responsive muscle; house mouse.		
SOURCE			
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; 1 (bases 1 to 9876)		
AUTHORS	Kemp,T.J., Sadusky,T.J., Saltisi,F., Carey,N., Moss,J., Yang,S., Sassoon,D.A., Goldspink,G. and Coulton,G.R.		
TITLE	Identification of Ankrd2, a novel skeletal muscle gene coding stretch-responsive ankryrin-repeat protein		
JOURNAL	Genomics 66 (3), 229-241 (2000)		
MEDLINE	20334618		
REFERENCE	2 (bases 1 to 9876)		
AUTHORS	Kemp,T.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-SEP-1999), Kemp T.J., Molecular Pathology, Imperial College School of Medicine, SAF Building, Exhibition Road, South Kensington, London SW7 2AZ, UNITED KINGDOM		
FEATURES	Location/Qualifiers		
source	1..9876		
gene	/organism="Mus musculus"		
TATA_signal	/strain="129/SvevTACfBr 129/SvevTACfBr"		
exon	/db_xref="taxon:10090"		
mRNA	/sex="female"		
CDS	/tissue_type="spleen"		
intron	573..9733		
repeat_region	/gene="Ankrd2"		
	573..578		
	/gene="Ankrd2"		
	/note="TATA signal is muscle specific"		
	613..697		
	/gene="Ankrd2"		
	/number=1		
	Join(613..697,4478..4579,4712..4870,4964..5071,6566994..7092,8376..8474,8666..8764,9484..9733)		
	/gene="Ankrd2"		
	Join(623..697,4478..4579,4712..4870,4964..5071,656994..7092,8376..8474,8666..8764,9484..9630)		
	/gene="Ankrd2"		
	/note="stretch responsive muscle"		
	/codon_start=1		
	/product="ankryrin repeat domain 2"		
	/protein_id="CAB99432.1"		
	/db_xref="GI:9501362"		
	/translation="MEGEPAVQATTELTEORLAQEETEKLRRSPAGKLSMDI ERLGVSPALQVKVQSERVKTSLDLRREIIVDVGQNLIELRKKRQKRR QPPPEPESITGVNNEETFLKANVEGKMKVIDKADGSGADTCDEFRTAL GMEILEKLENGAIVDFQDRDICTAMHWACRGHLEVVRLQSRGADTVNRV PLFVAVRTGHVEIVEHFVLSGLDIDINAKQEGSDALHDVRLNRYKIILLL AKNLACKTPTVLQWQADTRHALEHPEPESEQNLERPGSGRETPQDIPQAO 698..4477		
	/gene="Ankrd2"		
	/number=1		
	1020..1067		
	/note="CA repeat motif is polymorphic. Differing in between strains 129/SvevTACfBr, C57BL/6 and C57BL/10 /rpt family="CA repeat" /rpt unit=1020..1021		



```
/rpt_type=TANDEM
4478. .4579
/gene="Ankrd2"
/number=2
4580. .4711
/gene="Ankrd2"
/number=2
4712. .4870
/gene="Ankrd2"
/number=3
4871. .4963
/gene="Ankrd2"
/number=3
4964. .5071
/gene="Ankrd2"
/number=4
5072. .6576
/gene="Ankrd2"
/number=4
6577. .6675
/gene="Ankrd2"
/number=5
6676. .6993
/gene="Ankrd2"
/number=5
6994. .7092
/gene="Ankrd2"
/number=6
7093. .8375
/gene="Ankrd2"
/number=6
8376. .8474
/gene="Ankrd2"
/number=7
8475. .8665
/gene="Ankrd2"
/number=7
8666. .8764
/gene="Ankrd2"
/number=8
8765. .9483
/gene="Ankrd2"
/number=8
9484. .9733
/gene="Ankrd2"
/number=9
9714. .9719
polyA_signal /gene="Ankrd2"
BASE COUNT 2511 a 2355 c 2614 g 2390 t 6 others
ORIGIN
```

```
Query Match 11.4%; Score 131.8; DB 10; Length 9876;
Best Local Similarity 62.7%; Pred. NO. 5.8e-16;
Matches 240; Conservative 0; Mismatches 137; Indels 6; Gaps 2;

QY 195 agggcgaagcgcgtgcgaagcgtccctggaccctgcgcgggagatcatcgatgtgg 254
Db 4710 AGGGCCAGAGGCGGTGCCAGAGATCCCTGGACCTGCGACGTGAGATCATTTGATGG 4769

QY 255 gcggatccagaacctcatcgactgcggaagaacgcaagcagacgaagcgggacgtc 314
Db 4770 GTGGGATCCAGAACCTCATTGAAC TGAGGAAAAAGCAAGCAGAAAAAGCGAGATGCC 4829

QY 315 tggccgcctcgcgatgagcgcgccccagagcccgagagatcactggccctgtggatgag 374
Db 4830 TGGCTGCAGCCAGGAGCCTCTCCAGAGCCAGAGGAGATC-GTAAGGCTCCTGATTTAG 4888

QY 375 agaccttcctgaagctgcgtggagggaataaaggtcattgagaagttcctggctg 434
Db 4889 ACAGATGTGGGAGGGGGGAGCTGGGAGGCCCGCGAAGCTCCCTTTATTCTTTATTGGCGG 4948

QY 435 acgggggggtcagccgacacgtgcgaccagttccgtcggacagcactgcaccgagctccc 494
```

```
Db 4949 ATGTTCTACCTGCAGAC-----CGGCCCTGTGAATGAGGAGACATTCTGAAAGCTGCAG 5003
QY 495 tggaaagccacatgaaatcctcctgagaagcttctagataaatgggcccactgtgaactccc 554
Db 5004 TGGAGGGGAAAATGAAAGTCATTGACAAAGTACCTGGCGGACGGAGGTTTCAGCGGACACCT 5063
QY 555 aggatcggctggactgcacagcc 577
Db 5064 GTGATGAGGTGATCCCCCGCGGCC 5086
```

Search completed: April 17, 2002, 10:07:55  
Job time: 7119 sec

---

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2002, 08:10:37 ; Search time 26.67 Seconds  
(without alignments)  
1804.409 Million cell updates/sec

Title: US-09-758-593A-1  
Perfect score: 1693  
Sequence: 1 MEDSEAVORATLIEQRLAQ.....NLEGPNDSGRETQPVPQA 329

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_17.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1693	100.0	333	4 Q9GZV1	Q9gzv1 homo sapien
2	1484.5	87.7	328	11 Q9WV06	Q9wv06 mus musculus
3	677	40.0	319	6 Q9TU71	Q9tu71 oryctolagus
4	669.5	39.5	319	11 Q921F0	Q921f0 rattus norv
5	667	39.4	319	11 Q55014	Q55014 mus musculus
6	666	39.3	319	11 Q9CR42	Q9cr42 mus musculus
7	658	38.9	319	4 Q15327	Q15327 homo sapien
8	505	29.8	102	4 Q9NQC9	Q9nqc9 homo sapien
9	276	16.3	1214	5 Q9VUX2	Q9vux2 drosophila
10	249.5	14.7	1327	4 Q95271	Q95271 homo sapien
11	248.5	14.7	307	6 Q9N043	Q9n043 macaca fasc.
12	247.5	14.6	543	10 Q9LOP7	Q9lop7 arabidopsis
13	247.5	14.6	1166	4 Q9H2K2	Q9h2k2 homo sapien
14	247.5	14.6	1265	4 Q9HAS4	Q9has4 homo sapien
15	244	14.4	1050	4 Q9H014	Q9h014 homo sapien
16	241.5	14.3	592	12 Q9J4Z4	Q9j4z4 fowlpox vir
17	241.5	14.3	1136	6 Q9N180	Q9n180 bos taurus
18	240.5	14.2	548	10 Q9LVG7	Q9lvvg7 arabidopsis
19	239.5	14.1	1549	5 Q9V4B1	Q9v4b1 drosophila

20	238.5	14.1	1549	5	Q24241	Q24241 drosophila
21	238.5	14.1	1715	11	Q9ERD4	Q9erd4 rattus norv
22	238.5	14.1	1763	11	Q9EQG6	Q9eqg6 rattus norv
23	237.5	14.0	247	10	Q9ZQ79	Q9zq79 arabidopsis
24	235	13.9	1031	4	Q9UF42	Q9uf42 homo sapien
25	234.5	13.9	1719	4	Q13768	Q13768 homo sapien
26	234.5	13.9	1856	4	Q99407	Q99407 homo sapien
27	234	13.8	1159	5	Q9NCP8	Q9ncp8 drosophila
28	234	13.8	1430	11	Q9UJP7	Q9ujp7 mus musculus
29	234	13.8	2443	5	Q9VSA2	Q9vsa2 drosophila
30	233.5	13.8	1848	11	Q61302	Q61302 mus musculus
31	231.5	13.7	843	11	P97582	P97582 rattus norv
32	231	13.6	271	4	Q9NPL9	Q9npl9 homo sapien
33	231	13.6	1181	5	Q9XZ37	Q9xz37 drosophila
34	231	13.6	1181	5	Q9VBP3	Q9vbp3 drosophila
35	230	13.6	411	11	Q9D4D8	Q9d4d8 mus musculus
36	230	13.6	1943	11	Q61307	Q61307 mus musculus
37	229	13.5	4377	4	Q12955	Q12955 homo sapien
38	228.5	13.5	318	13	Q91974	Q91974 gallus gall
39	228.5	13.5	2119	5	Q9VAU5	Q9vau5 drosophila
40	228	13.5	637	5	Q9VUW9	Q9vuw9 drosophila
41	228	13.5	762	5	Q9VVR3	Q9vvr3 drosophila
42	228	13.5	1088	4	Q13484	Q13484 homo sapien
43	226.5	13.4	272	10	Q9AUW0	Q9auw0 oryza sativ
44	226	13.3	627	10	Q9SYK5	Q9syk5 arabidopsis
45	226	13.3	669	4	Q9BSK4	Q9bsk4 homo sapien

ALIGNMENTS

RESULT 1  
Q9GZV1 ID Q9GZV1 PRELIMINARY; PRT; 333 AA.  
AC Q9C.V1;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE SKELETAL MUSCLE ANKYRIN PROTEIN 2 (SKELETAL MUSCLE ANKYRIN REPEAT).  
GS ANKRD2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pallavicini A., Kojic S., Bean C., Vainzof M., Salamon M.,  
RA Ivoletta C., Bortoletto G., Pacchioni B., Trevisan S., Faulkner G.,  
RA Lanfranchi G., Valle G.;  
RL "Characterization of human skeletal muscle Ankr2.";  
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ304805; CAC19412.1;  
DR EMBL; AJ304804; CAC19411.1;  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 4.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 5.  
DR PROSITE; PS50088; ANK\_REPEAT; 4.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
SQ SEQUENCE 333 AA; 37151 MW; 679736F0491467A8 CRC64;

Query Match 100.0%; Score 1693; DB 4; Length 333;  
Best Local Similarity 100.0%; Pred. No. 3e-109;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MEDSEAVORATLIEQRLAQEEENEKLRGDTQKLPMDLLVLEDEKHHGSAALQKVKG	60
Db	5	MEDSEAVORATLIEQRLAQEEENEKLRGDTQKLPMDLLVLEDEKHHGSAALQKVKG	64
Qy	61	QERVKRTSLDLRREIIVGGIQLNIELKRRKKRKKRALLAASHPEPEITGPVDEET	120
Db	65	QERVKRTSLDLRREIIVGGIQLNIELKRRKKRKKRALLAASHPEPEITGPVDEET	124

Thu Apr 18 10:26:04 2002

```

121 FLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQD 180
125 FLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQD 184
181 RLDCATAMHWACRGHLEVVVKLLQSHGADTNVRDKLLSTPLHVAVRTGOVEIVEHFLSLGL 240
185 RLDCATAMHWACRGHLEVVVKLLQSHGADTNVRDKLLSTPLHVAVRTGOVEIVEHFLSLGL 244
241 EINARDREGDTALHDVAVRLNRYKIILLLHGADMTKNLAGTPTDLVOLWQADTRHAL 300
245 EINARDREGDTALHDVAVRLNRYKIILLLHGADMTKNLAGTPTDLVOLWQADTRHAL 304
301 EHPEPGAENHLEGNDSGRETPQVPAQ 329
305 EHPEPGAENHLEGNDSGRETPQVPAQ 333

```

```

RESULT 2
Q9WV06 PRELIMINARY; PRT; 328 AA.
AC Q9WV06;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SKELETAL MUSCLE AND CARDIAC PROTEIN (ANKYRIN REPEAT DOMAIN 2).
GN ANKRD2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP TISSUE-DIAPHRAGM MUSCLE;
RC Revolletia C., Formentin E., Lanfranchi G.;
RT "Characterization of a member of a new family proteins with ankyrin repeats."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VARIOUS STRAINS; TISSUE-SPLEEN, AND SKELETAL MUSCLE;
RX MEDLINE-20334618; PubMed-10873377;
RA Kemp T.J., Sadusky T.J., Saltisi F., Carey N., Moss J., Yang S.Y.,
RA Sassoon D.A., Goldspink G., Coulton G.R.;
RT "Identification of Ankr2, a novel skeletal muscle gene coding for a stretch responsive ankyrin-repeat protein."
RL Genomics 66:229-241(2000).
DR EMBL; AJ011118; CAB46546.1;
DR EMBL; AJ249346; CAB99432.1;
DR EMBL; AJ245514; CAB99431.1;
DR HSSP; P25963; I1KN.
DR MGD; MGI:1861447; Ankrd2.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 4.
DR SMART; SM00248; ANK; 4.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
SQ SEQUENCE 328 AA; 36707 MW; DB90D955EE9D175E CRC64;

```

```

Query Match 87.7%; Score 1484.5; DB 11; Length 328;
Best Local Similarity 88.8%; Pred. No. 7.1e-95;
Matches 292; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 1 MEDSAVORATALIEQRLAQQEENKLGRTFRQKLPMDLVLEDEKHHGAQSAALQKVK 60
DB 1 MEGPAVORATALIEQRLAQQEENKLGRTFRQKLPMDLVLEDEKHHGAQSAALQKVK 60
QY 61 QERVKRTSLDLRREIDYGGIONLIELKRRKQKRDALAAASHPEEITGPVDEET 120
DB 61 QERVKRTSLDLRREIDYGGIONLIELKRRKQKRDALAAASHPEEITGPVDEET 120
QY 121 FLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQD 180

```

```

Query Match 40.0%; Score 677; DB 6; Length 319;
Best Local Similarity 46.4%; Pred. No. 3.2e-39;
Matches 136; Conservative 54; Mismatches 69; Indels 34; Gaps 3;

QY 29 GDTROKLPMDLVLEDEKHHGAQSAALQKVKQGE-----RV 64
DB 18 GETGEFLP-----DDFRDQYEAATSEKQEDKLTPLAHVSLAEQWQWEREKQLEAE 70
QY 65 RKTSLDLRREIDYGGIONLIELKRRKQKRDALAAASHPEEITGPVDEETFLKA 124
DB 71 KKKLEQSRKLENLEDEEIIQLKKRKYRKTKVPVAK-----PEPELITEPVDVPFLKA 127
QY 125 AVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQDLDC 184
DB 128 ALENKLVAVKFLSDNNPDVCDYCKRTALHRACLEGLAIVEKLMGAGQIEFRDML 187
QY 185 TAMHWACRGHLEVVVKLLQSHGADTNVRDKLLSTPLHVAVRTGOVEIVEHFLSLGLEINA 244
DB 188 TAIHWACRGHLEVVVKLLQSHGADTNVRDKLLSTPLHVAVRTGOVEIVEHFLSLGLEINA 247
QY 245 RREGDTALHDVAVRLNRYKIILLLHGADMTKNLAGTPTDLVOLWQADTR 297
DB 248 KOREGDTPLHDVAVRLNRYKIILLLHGADMTKNLAGTPTDLVOLWQADTR 300

```

```

RESULT 4
Q9Z1F0 PRELIMINARY; PRT; 319 AA.
ID Q9Z1F0

```



Thu Apr 18 10:26:04 2002

Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
 Suroki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK009959; BAB26611.1; -;  
 DR EMBL: AK009655; BAB26419.1; -;  
 DR MGD: MGI:109621; ALP; ANK.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam: PF00023; ank; 4.  
 DR SMART: SM00248; ANK; 4.  
 DR PROSITE: PS50088; ANK\_REPEAT; 4.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 SQ SEQUENCE 319 AA; 36003 MW; 26A3C4062CF0E7D0 CRC64;

Query Match 39.3%; Score 656; DB 11; Length 319;  
 Best Local Similarity 47.7%; Pred. No. 1.8e-38;  
 Matches 132; Conservative 54; Mismatches 77; Indels 14; Gaps 3;

QY 45 EKHGGAQAALQKVK-GQER-----VRKTSLDRLREIIVGGIQLNLELRKKRQ 93  
 Db 40 EKQEDLTLPANSVKQGEQKSEKLEAEKLEKLEQSKLENLEFLEIIVOLKRRKY 99  
 QY 94 KKRDLAASHEPPPEITGPVDEETFLKAAVEGKMKVIEKFLADGGSDTCDOFRRTA 153  
 Db 100 KTKVPVVKVE---PEPIMTEPVDVPRFLAALNKLPPVEKFLSKNSPDVCDEYKRTA 156  
 QY 154 LHRASLEGHMEILKLDNGATVDFQDRDLCTAMHWACRGHLEVYKLLQSHGADTNVRD 213  
 Db 157 LHRACLEGLAIVEKLEAGAEIIFRDLMESTALHWACRGADNVLKLLNKGAKISARD 216  
 QY 214 KLLSLPLHVAVTGVEIVEHFLSLGLEINARDREGDTALHDVRLNRYKIIKLLHGA 273  
 Db 217 KLLSLHVAVTGVEIVEHFLSLGLEINARDREGDTALHDVRLNRYKIIKLLHGA 276  
 QY 274 DMATNLAGKPTDILVOLWQADTRHLEHPEPGAHN 310  
 Db 277 DLKVNKAGKTPMDLVLHWQSGTAFDLSLKENAYKN 313

RESULT 7  
 ID Q15327 PRELIMINARY; PRT; 319 AA.  
 AC Q15327;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE NUCLEAR PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SKIN;  
 RX MEDLINE=95247734; PubMed=7730328;  
 RA Chu W., Burns D.K., Swerlick R.A., Presky D.H.;  
 RA "Identification and characterization of a novel cytokine-inducible  
 RT nuclear protein from human endothelial cells.";  
 RL J. Biol. Chem. 270:10236-10245(1995).  
 DR EMBL: X83703; CAA58676.1; -;  
 DR HSSP: Q00421; IAWC.  
 DR InterPro: IPR002110; ANK.  
 DR Pfam: PF00023; ank; 4.  
 DR SMART: SM00248; ANK; 4.  
 DR PROSITE: PS50088; ANK\_REPEAT; 4.  
 DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
 KW Nuclear protein.  
 SQ SEQUENCE 319 AA; 36221 MW; 9C6E1D2B1C14A235 CRC64;

Query Match 39.9%; Score 658; DB 4; Length 319;  
 Best Local Similarity 51.8%; Pred. No. 6.5e-38;  
 Matches 129; Conservative 50; Mismatches 66; Indels 4; Gaps 2;  
 QY 49 GAQSAALQKVKGOERVKTSLDRLREIIVGGIQLNLELRKKRQKRDALAASHEPPPE 108  
 Db 56 GEQQWSEKQREAE-LPEKKLEQSKLENLEDEIIOLKKKKYKTKVPVYKE---PE 111  
 QY 109 PEEITGPVDEETFLKAAVEGKMKVIEKFLADGGSDTCDOFRRTALHRSLEGHMEILEK 168  
 Db 112 PEIITEPVDVPTFLKAANDENKLPVVEKFLSKNPNPDVCDEYKRTALHRSLEGHMEILEK 171  
 QY 169 LLDNGATVDFQDRDLCTAMHWACRGHLEVYKLLQSHGADTNVRDCLLSTPLHVAVTGQ 228  
 Db 172 LMEAGAQIEFRDLMESTALHWASRGNDLVKLLNKGAKISARDKLLSTALHVAVTGH 231  
 QY 229 VEIVEHFLSLGLEINARDREGDTALHDVRLNRYKIIKLLHGAADMMTKNLAGKTPD 288  
 Db 232 YECAEHLIACEADLNARDREGDTPLHDVRLNRYKIIKLLHGAADMMTKNLAGKTPD 291  
 QY 289 VOLWQADTR 297  
 Db 292 VLHWQNGTK 300

RESULT 8  
 ID Q9NQC9 PRELIMINARY; PRT; 102 AA.  
 AC Q9NQC9;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE ANKYRIN REPEAT DOMAIN 2 (FRAGMENT).  
 GN ANKRD2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SKELETAL MUSCLE;  
 RX MEDLINE=10873377;  
 RA Kemp T.J., Sadusky T.G., Saltisi F., Carey N., Moss J., Yang S.Y.,  
 RA Sassoon D.A., Goldspink G., Coulton G.R.;  
 RA "Identification of Ankrd2, a novel skeletal muscle gene coding for a  
 RT stretch-responsive ankyrin-repeat protein.";  
 RL Genomics 66:229-241(2000).  
 DR EMBL: AJ249975; CAB95416.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 102  
 FT NON\_TER 102  
 SQ SEQUENCE 102 AA; 11633 MW; 8DFD34DCB44D1E5 CRC64;  
 Query Match 29.8%; Score 505; DB 4; Length 102;  
 Best Local Similarity 99.0%; Pred. No. 5.4e-28;  
 Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 19 AOEENEKLRGDTROKLPMDLVLEDEKHGGAQAALQKVKGOERVKTSLDRLREIIV 78  
 Db 2 AOEENEKLRGDTROKLPMDLVLEDEKHGGAQAALQKVKGOERVKTSLDRLREIIV 61  
 QY 79 GGTONLEIKKKRQKRDALAASHEPPPEITGPVDEE 119  
 Db 62 GGIONLEIKKKRQKRDALAASHEPPPEITGPVDEE 102

RESULT 9  
 ID Q9VUX2 PRELIMINARY; PRT; 1214 AA.  
 AC Q9VUX2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)



QY 195 HLEVVKLQSHGADTNVRDKLSTPLHVAVRTGOVEIVEHFLSLGLEINARDREGDTALH 254  
DB 696 RVSVVEYLHGHGADHAKDGLPLHNAACSGYHVEAELLVRHGASVNVADLWKFTPLH 755  
QY 255 DAVLRNRYKIILKLLHGGADMMTNLAGTKTPTDLVOLWQADTRALE 301  
DB 756 EAAAGKYEICKLKLHKGADPTKKNRDNTPDLVKEGDTDIQDLK 802

RESULT 11  
Q9N043  
ID Q9N043 PRELIMINARY; PRT: 307 AA.  
AC Q9N043  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE UNNAMED PROTEIN PRODUCT.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,  
SU Suzuki Y., Sugano S., Hashimoto K.;  
RT "Isolation of full-length cDNA clones from macaque brain cDNA  
libraries."  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB046089; BAB01671.1;  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 7.  
DR SMART; SM00248; ANK; 6.  
DR PROSITE; PS50088; ANK\_REPEAT; 6.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
SQ SEQUENCE 307 AA; 32684 MW; D8E6EF1994504B95 CRC64;

Query Match 14.7%; Score 248.5; DB 6; Length 307;  
Best Local Similarity 31.0%; Pred. No. 1.1e-09;  
Matches 61; Conservative 32; Mismatches 71; Indels 33; Gaps 1;

QY 124 AAVEGKMKVIEKFLADGGSADTCDFRTRALHRSASLEHMEILEKLDNGATVDQDR-- 181  
DB 49 AANKGHLPPVQILKAGCDLVQDDGDTALHRAVTVGVNTEIIAALIHGCGALDRDRG 108  
QY 182 -----LDCTAMHWACRGHLEVVKLQSHGADTN 210  
DB 109 NTALEASWHGFSQSAKLLVRKAGANVLAKNAGNTALHACQNSHOSSTRVLLLAGSRAD 168  
QY 211 VRDKLLSTPLHVAVRTGOVEIVEHFLSLGLEINARDREGDTALHDAVRLNRYKIILKLL 270  
DB 169 LKNNAGDTCCLHVARYNHLISIRLLLSAFCSVHEKQAGDTALHAAALNKKVAKILLE 228  
QY 271 HGADMMTNLAGTKTPTD 287  
DB 229 AGADTTIVNAGOTPLE 245

RESULT 12  
Q9LQF7  
ID Q9LQF7 PRELIMINARY; PRT: 543 AA.  
AC Q9LQF7  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE F24B9.19 PROTEIN.  
GN F24B9.19.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Liu S., Yu G., Sakano H., Jhaveri A., Lee J., Lenz C., Pham P.,  
RA Toriumi M., Chin C., Choi E., Chou J., Gonzalez A., Chung M.,  
RA Howing B., Koo T., Li J., Liu A., Vaysberg M., Altafi H., Brooks S.,  
RA Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.,  
RA Johnson-Hopson C., Khan S., Kim C., Lam B., Nguyen M., Palm C.,  
RA Shian P., Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
RT "The sequence of BAC F24B9 from Arabidopsis thaliana chromosome 1."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Theologis A.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Theologis A.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. COLUMBIA;  
RA Theologis;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC007583; AAF7503.1;  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 8.  
DR SMART; SM00248; ANK; 6.  
DR PROSITE; PS50088; ANK\_REPEAT; 6.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
SQ SEQUENCE 543 AA; 58911 MW; C06C4C1B9E271DC8 CRC64;

Query Match 14.6%; Score 247.5; DB 10; Length 543;  
Best Local Similarity 29.1%; Pred. No. 2.6e-09;  
Matches 102; Conservative 48; Mismatches 121; Indels 79; Gaps 15;

QY 21 EEENEKLRGDTROKLPK-----DLLVLEDEKHHG-----AQSAAALQKVGK- 60  
DB 2 EGEEDTVAGSSIPKTKMMKQUTGKRDTLLHSAVRHGNKDRVETLTKTRESLNQLLKG 61  
QY 61 QERVKTSIDLRREITIDYGGION-----LIELKRKKOKKRDALAAASHEPPPEPEI 112  
DB 62 QNQSGETALYVAEYGVDEIVKEMINCYDLALVEI-----KARNGFDAFH----- 106  
QY 113 TGFVDEETFLKAABEGKMKVIEKFLADGGS--ADTCDFRTRALHRSASLEHMEILEKLL 170  
DB 107 -----IAAKQGDLDVL-KVLAESAHSALAMTVDLNSTTALHTAATQGTETVNVFLL 155  
QY 171 DNGATV-DFQDRIDCTAMHWACRGHLEVVK-LLOSHGADTNVRDKLSTPLHVAVRTGO 228  
DB 156 ELGSSAGTAGAKNGKTAHLSASRNGHVYKALLASEPAIAIRMDKKGQTALHMAVKGTN 215  
QY 229 VEIVEHFLSLG-LEINARDREGDTALHDAVRLNRYKIILKLLHG-ADMMTNLAGTKTPT 286  
DB 216 VEVVEELIKADRSSINATDKGNTALHIAARGRSOIVKLLANNMTDTKAVNRSGE-- 273  
QY 287 DLVQLWQADTRHALEHPE-----FGAEHNGLEGPNDSGRETPQPV 326  
DB 274 -----ALDTAEKIGNPEVALIILQKHGVSATKIKPSGNPA-RELKQTV 316

RESULT 13  
Q9H2K2  
ID Q9H2K2 PRELIMINARY; PRT: 1166 AA.  
AC Q9H2K2  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE TANKYRASE-LIKE PROTEIN (TANKYRASE 2).



GN TNKL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
FN [1]  
RN NCBI\_TaxID=9606;  
RP SEQUENCE FROM N.A.  
RC TISSUE-BREAST CARCINOMA;  
RA Kaimov A.N., Kuprash D.V., Petrov V.N., Vdovichenko K.K.,  
RA Bolichenko V.E., Scanlan M.J., Jongeneel C.V., Nedospasov S.A.,  
RA Lagarkova M.A.;  
RT "Cloning and characterization of TNKL, a member of tankyrase/ankyrin  
RT gene family.";  
RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kaminker P.G., Campisi J., Kim S.H., Yaswen P., Morin G.;  
RT "Rapid induction of apoptosis by ectopic expression of TANK2, a novel  
RT telomere-associated PARP.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Lyons R.J., Deane R., Lynch D.K., Ye Z.-S.J., Sanderson G.M.,  
RA Eyre H.J., Sutherland G.R., Daly R.J.;  
RT "Identification of a novel human tankyrase through its interaction  
RT with the adaptor protein Grb14.";  
RL J. Biol. Chem. 0:0-0(2001).  
DR EMBL: AF264912; AAK44694.1; -  
DR EMBL: AF342982; AAK25811.1; -  
DR EMBL: AF329696; AAK13463.1; -  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000345; CytC\_heme\_bind.  
DR InterPro: IPR001660; SAM.  
DR Pfam: PF00023; ank; 16.  
DR Pfam: PF00536; SAM; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00454; SAM; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 15.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
DR NON\_TER 1  
SQ SEQUENCE 1166 AA; 126917 MW; 4C8B38BD97CE704 CRC64;

Query Match 14.6%; Score 247.5; DB 4; Length 1166;  
Best Local Similarity 23.4%; Pred. No. 6.7e-09;  
Matches 105; Conservative 49; Mismatches 133; Indels 161; Gaps 13;  
Qy 34 NEKLGRDTRQKLPMDLLVLEDEKHHGAQSAALQKVGQERVKTS----LDLRRRIIDVG 79  
Db 408 SEKAND-----VVEVVVKEAKVNDNL-GOTSLHRAAYCGHLQTCRLLSYG 456  
Qy 80 GTONLIELR-----KRRQKRRDALA-ASHEPPPEPEITGPVDEETFLK----- 123  
Db 457 VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVPLHNAACSGYHVAE 516  
Qy 124 -----AAVEGKMKVIERFLADGGSDATCDQFRRTALHRASLEGHMEILE 167  
Db 517 VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVPLHNAACSGYHVAE 576  
Qy 168 KLLDNGATVDFQDRDCTAMHWACRGHLEVVKLLQSHGA-----D 208  
Db 577 LLVKGAVNVADLWKFTPLHEAAKGYETCKLLQHGADPTKKNRGDNTPLDLVKDGD 636  
Qy 209 TNVRDKLL-----STPLHVAVRTGOVEIVE 233  
Db 637 TDIOQLLRGDAALLDAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAE 696  
Qy 234 HFLSLGLEINARDREG-----DTALHDAVRLN 260  
Db 697 YLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHEAAQKG 756  
Qy 261 RYKIKLKLHLLHGADMTNLAGKTPTDLVQWQADTRHAL-----EHPE--- 304

Db 757 RTQICALLAHGADPTLKNQEGQTPDLV-----SADDVSALLTAAMPSPALPSYKQPVLN 813  
Qy 305 ----PGAEHNGLEGPNDSGRETPQPVPA 328  
Db 814 GVRSPGATADAL-----SSGSPSSPSLSA 837  
RESULT 14  
Q9HAS4 PRELIMINARY; PRT; 1265 AA.  
ID Q9HAS4  
AC Q9HAS4  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE TANKYRASE-RELATED PROTEIN (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
FN [1]  
RN SEQUENCE FROM N.A.  
RP Mor.: D.W., Meese E.;  
RA "Novel tankyrase-related gene.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF305081; AAG25674.1; -  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000345; CytC\_heme\_bind.  
DR InterPro: IPR001660; SAM.  
DR Pfam: PF00023; ank; 16.  
DR Pfam: PF00536; SAM; 1.  
DR PRINTS: PR01415; ANKYRIN.  
DR SMART: SM00248; ANK; 15.  
DR SMART: SM00454; SAM; 1.  
DR PROSITE: PS50088; ANK\_REPEAT; 15.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
DR NON\_TER 1  
SQ SEQUENCE 1265 AA; 136905 MW; 3BFD7DE7AEBC038 CRC64;

Query Match 14.6%; Score 247.5; DB 4; Length 1265;  
Best Local Similarity 23.4%; Pred. No. 7.5e-09;  
Matches 105; Conservative 49; Mismatches 133; Indels 161; Gaps 13;  
Qy 24 NEKLGRDTRQKLPMDLLVLEDEKHHGAQSAALQKVGQERVKTS----LDLRRRIIDVG 79  
Db 507 SEKAND-----VVEVVVKEAKVNDNL-GOTSLHRAAYCGHLQTCRLLSYG 555  
Qy 80 GTONLIELR-----KRRQKRRDALA-ASHEPPPEPEITGPVDEETFLK----- 123  
Db 556 CDPNIISLQGTALQMGNNVQQLQEGISLGNSEADRLLEAAKAGDVETVKLCTVQS 615  
Qy 124 -----AAVEGKMKVIERFLADGGSDATCDQFRRTALHRASLEGHMEILE 167  
Db 616 VNCRDIEGRQSTPLHFAAGYNRVSVVEYLLQHGADVHAKDGGVPLHNAACSGYHVAE 675  
Qy 168 KLLDNGATVDFQDRDCTAMHWACRGHLEVVKLLQSHGA-----D 208  
Db 676 LLVKGAVNVADLWKFTPLHEAAKGYETCKLLQHGADPTKKNRGDNTPLDLVKDGD 735  
Qy 209 TNVRDKLL-----STPLHVAVRTGOVEIVE 233  
Db 736 TDIOQLLRGDAALLDAKKGCLARVKKLSSPDNVNCRDTQGRHSTPLHLAAGYNNLEVAE 795  
Qy 234 HFLSLGLEINARDREG-----DTALHDAVRLN 260  
Db 796 YLLQHGADVNAQDKGLIPLHNAASYGHVDVAALLIKYNACVNAATDKWAFPLHEAAQKG 855  
Qy 261 RYKIKLKLHLLHGADMTNLAGKTPTDLVQWQADTRHAL-----EHPE--- 304  
Db 856 RTQICALLAHGADPTLKNQEGQTPDLV-----SADDVSALLTAAMPSPALPSYKQPVLN 912

```
QY 305 ----PGAENHGLEGNDSGRETPQVPA 328
Db 913 GVRSPGATADAL-----SSGPSSPSSLSA 936

RESULT 15
Q9H014 PRELIMINARY; PRT; 1050 AA.
AC Q9H014;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE HYPOTHETICAL 117.0 KDA PROTEIN.
GN DKFZP434L0718.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL136784; CAB66718.1;
DR InterPro; IPR002110; ANK.
DR InterPro; IPR003123; VPS9.
DR Pfam; PF00023; ank; 8.
DR Pfam; PF02204; VPS9; 1.
DR PRINTS; PR01415; ANKYRIN.
DR SMART; SM00248; ANK; 8.
DR SMART; SM00167; VPS9; 1.
DR PROSITE; PS00088; ANK_REPEAT; 8.
DR PROSITE; PS0297; ANK_REPEAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 1050 AA; 117012 MW; 6A0D4F77F3C1D11D CRC64;

Query Match 14.4%; Score 244; DB 4; Length 1050;
Best Local Similarity 25.6%; Pred. No. 1e-08;
Matches 87; Conservative 46; Mismatches 117; Indels 90; Gaps 10;

QY 24 NEKLRGDTROKLPMDLL-----VLEDEKHGAQAALQKVGQERVKTSLD--LRRE 74
Db 562 NEK--GDT----PLHIAARWGYQGVETLLQNGASTEI-----QNRLKETPLKCALNSK 609
QY 75 IIDVGGTQNLIELRKKRKKRDLAASHEPPPEPEIT-----GPVDEET- 120
Db 610 ILSV---WEAYHLSFERRQKSSAPVQS--PQRSVDSISQESSTSFSSMSAGSRQETK 664
QY 121 -----FLKAAVEGKMKVIEKFL----- 137
Db 665 KDYREVEKLLRAVDGDLNMYRYLLEWTEDEDAEDTVSAADPEFCHPLCQCPKCAPAQ 724
Y 138 -----ADGSADTCQFRTALHRASLEGHMEILEKLLDNGATVDFODRLDCTAMHWA 190
Db 725 KRLAKVPASGLGVNVTSDGSSPLHVAALHGRADLIRLLKLGANAGARNADQAVPLHLA 784
QY 191 CRGGHLEVVKLQSHGADTNVRDKLLSTPLHVAVRTQGVIEVHFLSLGLEINARDREGD 250
Db 785 COGHFQVVKLLDSNAKPNKDLGNTPLIYACSGGHHVALLLQHGASINASNKN 844
QY 251 TALHDAVRLNRYKIILKLLHGDMMTKNLAKGTPDVLQ 290
Db 845 TALHEAVIEKHVFVVELLLHGHASVQVNLNKRQRTAVDCAE 884
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2002, 08:09:42 ; Search time 23.95 Seconds  
(without alignments)  
1017.541 Million cell updates/sec

Title: US-09-758-593a-1

Perfect score: 1693

Sequence: 1 MEDSEAVQATALIEQRLAQ.....NGLEGPNDSGRETPQVPVPAQ 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_1101.\*

1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1688	99.7	329	22	Human polypeptide
2	892	52.7	182	22	Human polypeptide
3	658	38.9	319	22	Human shear stress
4	296.5	17.5	510	22	Human polypeptide
5	282.5	16.7	978	21	Human OREF ORF2052
6	249.5	14.7	949	21	Human truncated ta
7	249.5	14.7	1327	21	Human tankyrase I
8	249.5	14.7	1327	21	Human tankyrase
9	249.5	14.7	1327	22	Human tankyrase I
10	248	14.6	1166	22	Human SPANK. Homo
11	247.5	14.6	784	22	Human tankyrase2 c

12	247.5	14.6	1074	20	Human Grb7 effecto
13	247.5	14.6	1100	22	Tankyrase homology
14	247.5	14.6	1166	21	Human tankyrase II
15	247.5	14.6	1166	22	Human tankyrase2 T
16	247.5	14.6	1169	22	Human tankyrase2 I
17	247.5	14.6	1169	22	Human tankyrase2 c
18	247.5	14.6	1240	22	Tankyrase homology
19	247.5	14.6	1262	22	Human tankyrase2 c
20	247.5	14.6	1385	22	Human tankyrase2 T
21	246.5	14.6	756	22	Human tankyrase2 c
22	246	14.5	991	22	Mouse SPANK. Mus
23	245.5	14.5	907	22	Human breast cance
24	244	14.4	1166	22	Human tankyrase ho
25	240.5	14.2	868	21	Arabidopsis thalia
26	240.5	14.2	879	21	Arabidopsis thalia
27	240.5	14.2	890	21	Arabidopsis thalia
28	237.5	14.0	522	22	Human tankyrase2 c
29	235	13.9	551	22	Human death domain
30	235	13.9	1715	22	Human polypeptide
31	235	13.9	1715	22	Human polypeptide
32	233	13.8	765	22	Human protein kina
33	231	13.6	1181	22	Drosophila tankyra
34	230.5	13.6	362	22	Human TRF1 TANK2 b
35	230.5	13.6	840	22	Human polypeptide
36	230	13.6	705	22	Human colon cancer
37	228	13.5	1088	22	Human kidney ankyr
38	228	13.5	1333	21	Human tankyrase II
39	224	13.2	679	15	2-5A-dependent RNA
40	224	13.2	679	16	Partial murine 2-5
41	224	13.2	679	18	Mouse 2-5A-depende
42	224	13.2	679	20	Mouse 2-5A-depende
43	223	13.2	422	15	Fowlpox virus (FPV
44	223	13.2	422	15	Fowlpox virus prot
45	223	13.2	422	18	Homology vector 44

#### ALIGNMENTS

RESULT 1  
ID AAM40353 standard; Protein; 329 AA.

AC AAM40353;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 3498.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
chemokine; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

PR 21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

OS Homo sapiens.  
XX WO270153312-A1.  
XX 26-JUL-2001.  
XX 26-DEC-2000; 2000WO-US34263.  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.  
XX 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
XX N-PSDB; AAI61295.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -  
Example 2; SEQ ID NO 7070; 10078pp; English.  
The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Query Match 52.7%; Score 892; DB 22; Length 182;  
Best Local Similarity 99.4%; Pred. No. 1.1e-78;  
Matches 178; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MEDSEAVQATATLIEORLAQEEENKLRGDTROKLPMDLLVLEDEKHGGAQSAALQKVKG 60  
DB 4 MEDSEAVQATATLIEORLAQEEENKLRGDTROKLPMDLLVLEDEKHGGAQSAALQKVKG 63  
QY 61 QERVKTSLDLRREIIDVGGIQLNIELRRKKQKRRDALAASHEPPEEITGPVDEET 120  
DB 64 QERVKTSLDLRREIIDVGGIQLNIELRRKKQKRRDALAASHEPPEEITGPVDEET 123  
QY 121 FLKAAVEGKMKVIEKFLADGSGADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQ 179  
DB 124 FLKAAVEGKMKVIEKFLADGSGADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQ 182

(HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
XX N-PSDB; AAI59509.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -  
Example 6; SEQ ID NO 3498; 10078pp; English.  
The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Query Match 99.7%; Score 1688; DB 22; Length 329;  
Best Local Similarity 99.7%; Pred. No. 1.3e-155;  
Matches 328; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MEDSEAVQATATLIEORLAQEEENKLRGDTROKLPMDLLVLEDEKHGGAQSAALQKVKG 60  
DB 1 MEDSEAVQATATLIEORLAQEEENKLRGDTROKLPMDLLVLEDEKHGGAQSAALQKVKG 60  
QY 61 QERVKTSLDLRREIIDVGGIQLNIELRRKKQKRRDALAASHEPPEEITGPVDEET 120  
DB 61 QERVKTSLDLRREIIDVGGIQLNIELRRKKQKRRDALAASHEPPEEITGPVDEET 120  
QY 121 FLKAAVEGKMKVIEKFLADGSGADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQ 180  
DB 121 FLKAAVEGKMKVIEKFLADGSGADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQ 180  
QY 181 RLDCATAMHWACRGHLEVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIHFSLGL 240  
DB 181 RLDCATAMHWACRGHLEVYKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIHFSLGL 240  
QY 241 EINARDREGDTALHDVRLNRYKIILLHGHGDMTKNLGKTPDVLVOLWQADTRHAL 300  
DB 241 EINARDREGDTALHDVRLNRYKIILLHGHGDMTKNLGKTPDVLVOLWQADTRHAL 300  
QY 301 EHPEPGAENHGLEGNDSGRETPQVPAQ 329  
DB 301 EHPEPGAENHGLEGNDSGRETPQVPAQ 329

RESULT 2  
ID AAM42139  
XX AAM42139 standard; Protein; 182 AA.  
XX AC AAM42139;  
XX 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 7070.



Thu Apr 18 10:26:02 2002

124 AAVEGKMKVTEKFLADGSGADTCDFRRFALHRSLEHMEILEKLLDNGATVDFQDRLD 183  
 27 aaanavkcaevlplssvvnvsgrrtalhaaIngvnmvlllkganinafdkdkd 86  
 184 CTAMHACRGHLEVVLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGLEIN 243  
 87 rralhwaaymgldvallinhgaevtcckdkgytphaaasngvnnvkhllnlgveid 146  
 244 ARDEGDTALHDVRLNRYKIILKLLHGDGMTKLNAGTKPTDVLQWQADTRHAL 300  
 147 einvyngntalciacyngdavnvelidyanvnpnnngftp---lnfaaasthgal 200

RESULT 5  
 AAB42286  
 ID AAB42288 standard; Protein: 978 AA.  
 AC AAB42288;  
 DT 08-FEB-2001 (first entry)  
 DE Human ORFX ORF2052 polypeptide sequence SEQ ID NO:4104.

Human: open reading frame; ORFX: detection; cytotatic; hepatotropic;  
 vulnary; antiporiatic; antiparkinsonian; nootropic; neuroprotective;  
 anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 cholesterol ester storage; systemic lupus erythematosus; infection;  
 severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 bone damage; cartilage damage; antiinflammatory disease; coagulation;  
 thrombosis; contraceptive.

\*Homo sapiens.  
 WO200058473-A2.  
 05-OCT-2000.  
 31-MAR-2000; 2000WO-US08621.  
 31-MAR-1999; 99US-0127607.  
 02-APR-1999; 99US-0127636.  
 05-APR-1999; 99US-0127728.  
 30-MAR-2000; 2000US-0540763.  
 (CURA-) CURAGEN CORP.  
 Shimkets RA, Leach M;  
 WPI; 2000-602362/57.  
 N-PSDB; AAC76497.  
 Novel nucleic acids and peptides derived from open reading frame x,  
 useful for treating e.g. cancers, proliferative disorders,  
 neurodegenerative disorders and cardiovascular disease -  
 Claim 11; Page 3289-3292; 5507pp; English.

AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
 which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 sequences have activities such as: cytotatic; hepatotropic; vulnary;  
 antiporiatic; antiparkinsonian; nootropic; neuroprotective;  
 osteopathic; anticonvulsant; antithyroid; immunosuppressant;  
 immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 antidiabetic; hypotensive; dermatological; immunosuppressive;  
 antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
 antithyroid; and antianaemic. The sequences can be used for determining

the presence of or predisposition to, or preventing or treating  
 pathological conditions associated with an ORFX-associated disorder. The  
 nucleic acids can be used to express ORFX proteins in gene therapy  
 vectors. The proteins and nucleic acids may be used to treat cancers,  
 proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 graft vs host disease, cardiovascular disease, diabetes mellitus,  
 hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
 coagulation; to inhibit thrombosis; and as a contraceptive.

Sequence 978 AA;  
 Query Match 16.7%; Score 282.5; DB 21; Length 978;  
 Best Local Similarity 28.1%, Pred. No. 1.7e-18;  
 Matches 94; Conservative 44; Mismatches 120; Indels 77; Gaps 9;  
 QY 28 RGDTRKLPMDLLVLEDEK-----HHGAQSAALQKVGQERYVKTSLDLRREII 76  
 DB 347 qpqgketlesalialdsekpkklrfhpkqlfysarggelqkvl-----im 391  
 QY 77 DVGGIONLIEKRKKQKRDALAAASHEPPPEETGPDV-----aghdichmlvqaganidtcseq 440  
 DB 392 lvdgldpnf---kmeqknkrspshaae-----aghdichmlvqaganidtcseq 440  
 QY 118 EEFFLKAAVEGKMKVIEKFLADGSGADTCDFRRFALHRSLEHMEILEKLLDNG-ATV 176  
 DB 441 rtplmeaanhhleavkylkagaldvdpkdaegstchlhaakghyevvyllsgrmdv 500  
 QY 177 DFQDRLDCTAMHACRGHLEVVLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFL 236  
 DB 501 ncqddgwtpmiawatekghvdlvklillskgsdinirneeniclhwafsgcvdiaeill 560  
 QY 237 SLGLEINARDREGDTALHDVRLNRYKIILKLLHGDGMTKLNAGTKPTDVLV---QLW 292  
 DB 561 aakcdlhhavnihgdsphlaarenrydcvvlfsrdsdvtlknkegetplqcaslnsqw 620  
 QY 293 QA-DYTHALEHPEPGAENHGLEGNDSGRETPQPV 326  
 DB 621 salqmskalq-----dsaprpapv 640

RESULT 6  
 AAY44404  
 ID AAY44404 standard; protein: 949 AA.  
 XX AAY44404;  
 AC AAY44404;  
 DT 22-MAR-2000 (first entry)  
 DE Human truncated tankyrase-2.  
 XX Human tankyrase; TRF1; telomeric repeat binding factor-1;  
 KW telomere length regulation; recombinant tankyrase; aging; skin atrophy;  
 KW macular degeneration; atherosclerosis; ataxia telangiectasia;  
 KW tankyrase modulator; rational drug design.  
 OS Hemo sapiens.  
 XX WC9964606-A1.  
 PN 16-DEC-1999.  
 PD 09-JUN-1999; 99WO-US12968.  
 PF 10-JUN-1998; 98US-0095225.  
 PR 17-AUG-1998; 98US-0135233.  
 PR 19-NOV-1998; 98US-0196387.  
 XX (UYRQ ) UNIV ROCKEFELLER.



Thu Apr 18 10:26:02 2002

ataxia telangiectasia; tankyrase modulator; rational drug design.

Homo sapiens.

Key Location/Qualifiers

Domain 181..1010

note="ANKyrin-specific (ANK) repeat consensus domain"

Domain 1023..1086

/label= Sterile\_alpha\_motif

Domain 1176..1314

/note= "Poly(ADP-ribose)polymerase (PARP) related domain"

WO9964606-A1.

16-DEC-1999.

09-JUN-1999; 99WO-US12968.

10-JUN-1998; 98US-0095225.

17-AUG-1998; 98US-0135233.

19-NOV-1998; 98US-0196387.

(UYRQ ) UNIV ROCKEFELLER.

De Lange T, Smith S;

WPI: 2000-116549/10.

N-PSDB; AA229627.

New nucleic acid encoding vertebrate tankyrase, a regulator of telomere length, used to identify modulators, e.g. for inhibiting growth of cancer

- Claim 1; Fig 1B; 133pp; English.

The present sequence is human tankyrase, a TRF1 (telomeric repeat binding factor-1) binding protein. This was isolated using a two-hybrid screen with TRF1. Tankyrase is involved in regulation of telomere length through TRF1 and may directly modulate the effect of TRF1. Recombinant tankyrase, or its fragments are used to identify specific modulators which are potential drugs for countering telomere shortening associated with aging (e.g. atrophy of skin, macular degeneration or atherosclerosis) or diseases like ataxia telangiectasia. They are also used to raise specific antibodies (used as immunoassay reagents and as modulators of tankyrase activity) and in rational drug design.

Sequence 1327 AA;

Query Match 14.7%; Score 249.5; DB 21; Length 1327;

Best Local Similarity 26.2%; Pred. No. 4e-15;

Matches 91; Conservative 33; Mismatches 110; Indels 113; Gaps 9;

40 LVLEDEKH---HGAQSAALQKVGQERVKTSIDLRRLIIDVGGION-----LIELR 88

484 ltyefkghslldaareadlakkv-----kt---lalelnfkqpqshetahcavaslh 534

89 KKRKQ-----KKRDALAAASHEPPPEEITGPVDEETFLKAAVEGKMKVIE 134

535 pkrkqvteillrrkganvneknkdfmtplh-----vaaerahndvme 575

135 KFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNG----- 173

576 vlkhgkmaaldtlgqtaalhraalaghiqtrclllsygsdpsiiislgftaaqmgneav 635

174 -----ATVDFQ-----DRLDC-----TAMHWACRGG 194

636 qqilsestprtssdvryrllleaskagdlctvklcssqnvncrdleghrstplhfaagyn 695

195 HLEVYKLLQSHGADTNVRDKLLSTPLHVAVRTQVEIVHFSLGLEINARDREGDTALH 254

696 rvsveyllbhgadvhakdgkglvplhnacsyghyevaelvrrhgasvnnvadlwktfplh 755

255 DAVRLNRYKIILKLLLLHGDMMTKNLAKGTPTDIVQLWQADTRHALE 301

756 eaaakgyeickillllkngadptknrdgntpldivkegdtidqllk 802

RESULT 9

AA66279

ID AAB66279 standard; Protein; 1327 AA.

XX

AC AAB66279;

XX

DT 05-APR-2001 (first entry)

XX

DE Human tankyrase1 SEQ ID NO: 4.

XX

KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;

KW inflammatory disorder.

XX

OS Homo sapiens.

XX

PN WO200100849-A1.

XX

PD 04-JAN-2001.

XX

PF 28-JUN-2000; 2000WO-US17827.

XX

PR 29-JUN-1999; 99US-0141582.

XX

PA (ICOS-) ICOS CORP.

XX

PI Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX

DR WPI: 2001-102896/11.

XX

DR N-PSDB; AAF63838.

XX

PT New tankyrase2 polypeptides, useful for treating conditions mediated by

PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,

PT inflammatory and autoimmune disorders

XX

PS Example 1; Page 118-121; 242pp; English.

XX

CC The present invention provides the protein and coding sequence for the

CC human tankyrase2 protein. This is found in two different versions,

CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has

CC polyADP-ribosylation activity and is involved in the modification of

CC TRF1, which is a telomere-specific binding protein. The regulation of

CC telomere length, in which TRF1 has a role, is linked to ageing and

CC cancer. The sequences are useful in the treatment of cancers and

CC inflammatory disorders.

XX

Sequence 1327 AA;

Query Match 14.7%; Score 249.5; DB 22; Length 1327;

Best Local Similarity 26.2%; Pred. No. 4e-15;

Matches 91; Conservative 33; Mismatches 110; Indels 113; Gaps 9;

40 LVLEDEKH---HGAQSAALQKVGQERVKTSIDLRRLIIDVGGION-----LIELR 88

484 ltyefkghslldaareadlakkv-----kt---lalelnfkqpqshetahcavaslh 534

89 KKRKQ-----KKRDALAAASHEPPPEEITGPVDEETFLKAAVEGKMKVIE 134

535 pkrkqvteillrrkganvneknkdfmtplh-----vaaerahndvme 575

135 KFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNG----- 173

576 vlkhgkmaaldtlgqtaalhraalaghiqtrclllsygsdpsiiislgftaaqmgneav 635

174 -----ATVDFQ-----DRLDC-----TAMHWACRGG 194

636 qqilsestprtssdvryrllleaskagdlctvklcssqnvncrdleghrstplhfaagyn 695





Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
WPI; 2001-102896/11.  
N-PSDB; AAF63917.

New tankyrase2 polypeptides, useful for treating conditions mediated by poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers, inflammatory and autoimmune disorders -

Example 1; Page 144-146; 242pp; English.

The present invention provides the protein and coding sequence for the human tankyrase2 protein. This is found in two different versions, designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has polyADP-ribosylation activity and is involved in the modification of TRF1, which is a telomere-specific binding protein. The regulation of telomere length, in which TRF1 has a role, is linked to ageing and cancer. The sequences are useful in the treatment of cancers and inflammatory disorders.

Sequence        784 AA;

Query Match                  14.6%; Score 247.5; DB 22; Length 784;  
Best Local Similarity      23.4%, Pred. No. 3.le-15;  
Matches 105; Conservative 49; Mismatches 133; Indels 161; Gaps 13;

QY     24   NKLKRGDTQRQLPMDLILLEDEKHGASAAALQVKYGQERVKTS-----LDLREIIDVG    79  
Db     :|||    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |  
26   sekahnd-----vvevvvkheakvnaldnl-ggtslhraaycghlqtcrlillsyg    74  
:  
  
QY     80   GIONLIETLR-----KKRKOKRDALA-ASHPEPPPEPETTGPVDEETFLLK-----    123  
Db     ||||    :    :    :    :    :    :    :    :    :    :    :    :    :    :    :  
75   cdpniislgftalgmgneuvdqllqegislgnseadqrllaakaagdvetykklctvqs    134  
:  
  
QY     124 -----AAVEGKMKVIEKFADGGSSADTCDFRFTALHRASLEGHMEILE    167  
Db     ||::||    |    |    |    |    |    |    |    |    |    |    |    |    |    |  
135   vncrdiegrqstplhfaagynrvsvveyllhqgadvhakdkggvlphlnacsygyhevae    194  
:  
  
QY     168   KLIDNGATVDQRLDCTAMHWACRGHLEVVKLLQSHGA-----D    208  
Db     |:|:|||    :    :    :    :    :    :    :    :    :    :    :    :    :    :  
195   llvkhgavnavvadlwktfplheaakgkyeikllqhgdaptkknrgntpldlvkdgd    254  
:  
  
QY     209   TNVRDKILL-----STPLHAVRTGVGOEIVE    233  
Db     ||::::|    |    |    |    |    |    |    |    |    |    |    |    |    |  
255   tdiqlilrldaalldaakkglarvkvkllsspdnvncrtqtgrbstplhiaagynnlevae    314  
:  
  
QY     234   HFUSLGLEINARDREG-----DTALHDVAURLN    260  
Db     |    |    |    |    |    |    |    |    |    |    |    |    |    |  
315   yllhqgadavnraqkggilplhmaasyghvdvaalllkynacvnatdkwaftplheaackg    374  
:  
  
QY     261   RYKIILKILLHGADMMTKNLAGTKPTDLVLQIWAQDTHAL-----EHPE---    304  
Db     |    |    |    |    |    |    |    |    |    |    |    |    |    |  
375   rtqicalllahgaadtliknqeggtpldlv----saddvsalltaampspaispcykqpqln    431  
:  
  
QY     305 ----PGAENHGLEPNDSGRTEPQPVA    328  
Db     ||||    |    |    |    |    |    |    |    |    |    |    |    |    |  
432   gvrspgatadal----ssgpspspsisa    455  
:

RESULT 12  
AAY05734 standard; Protein; 1074 AA.  
XX AC AAY05734;  
XX DT 19-JUL-1999 (first entry)  
XX DE Human Grb7 effector 2.2412 protein.  
KW Grb7 effector; 2.2412 protein; human; signal transduction;  
KW tumour marker; breast cancer; prostate cancer; prognosis;  
KW diagnosis.

Db 605 yllqhgadnvaqdkggliplnhaasyghdvaaallikynaslnatdkwaftplheaackg 664  
Qy 261 RYKIILKLLHGGADMTKLNLAGKTPDLVQLWQADTRHAL-----EHPE--- 304  
Db 665 rtqlcallahgadtlnkgqgtpdlv---saddvsallitaampsalpsckypqvl 721  
Qy 305 ----PGAENHNGLEGNDSGRETPQVPVA 328  
Db 722 gvrspgatal-----ssgspsspslsa 745

RESULT 13  
AAY97748 standard; Protein; 1100 AA.  
AAY97748;  
06-AUG-2001 (first entry)  
Tankyrase homologue isotype 1 protein sequence.  
Tankyrase homologue isotype 1; TaHo-1; TaHo-2; cell proliferation;  
cell cycle protein; cell cycle associated disorder; cancer; gene mapping;  
chromosome mapping; gene therapy; vaccine.  
Unidentified.  
WO200130987-A2.  
03-MAY-2001.  
25-OCT-2000; 2000WO-US41528.  
25-OCT-1999; 99US-0427154.  
(RIGE-) RIGEL PHARM INC.  
Luo Y, Chan E, Xu X, Huang B;  
WPI: 2001-300503/31.  
N-PSDB; AAA91487.  
Novel recombinant cell cycle polypeptide, tankyrase H useful for  
inducing or preventing cell proliferation in cells, and for diagnosing,  
treating or preventing cell cycle associated disorders such as cancer  
Claim 22; Fig 3; 63pp; English.  
This sequence is the Tankyrase homologue isotype 1 (TaHo-1) protein  
of the invention. The invention also relates to the TaHo-2 protein.  
The TaHo proteins are useful for inducing or preventing cell  
proliferation in cells, and in the study or treatment of conditions  
mediated by the cell cycle proteins, such as to diagnose, treat or  
prevent cell cycle associated disorders, preferably cancer. The TaHo  
coding sequences are useful as hybridisation probes, in chromosome and  
gene mapping and in the generation of anti-sense DNA and RNA. The coding  
sequences are also useful for the preparation of TaHo, for generating  
either transgenic animals or knock out animals which, in turn, are useful  
in a development and screening of therapeutically useful agents, in gene  
therapy, as vaccine, and for construction of hybridisation probes for  
mapping the gene which encodes TaHo and for the genetic analysis of  
individuals with genetic disorders. The TaHo proteins, and their coding  
sequences are useful in screening assays.

Query Match 14.6%; Score 247.5; DB 22; Length 1100;  
Best Local Similarity 23.4%; Pred. No. 4.9e-15;  
Matches 105; Conservative 49; Mismatches 133; Indels 161; Gaps 13;  
Qy 24 NEKRGDTRQKLPMDLLVEDEKHHGAQSAALQKVGQERVRKTS-----LDLRRREIIDVG 79

Db 342 sekahnd-----vvevvvkheakvnaIdnl-gqtslhraaycghlgtcrlllsyg 390  
Qy 80 GIONLIELR-----KKRQKKRDALA-ASHEPPEPEPEITGPVDEETFLK----- 123  
Db 391 cdpnlislgfftalqmgngnvvqllqegislgneadrtqlleaakagdvettvkictvqs 450  
Qy 124 -----AAVEGKMKVIEFLADGGSDATCDQFRRTALHRALESLEGHMEILE 167  
Db 451 vncrdiegrqstplhfaagynrvsvveyllqhgadvhakdkgglvplhnacsyghyvae 510  
Qy 168 KLLNGATVDFQDLDCDTAMHWACRGHLEVVKLLQSHGA-----D 208  
Db 511 llvkhgavnnvadlwktplheaakgyeickllllqhgadtppdkknrdgntpldlvkdg 570  
Qy 209 TNVRDKLL-----STPLHVAVRTGOVEIVE 233  
Db 571 tdqdlirgdaalldaakkgclarvkkisspndvncrdtqgrhstplhlaagynnlevae 530  
Qy 234 HFLSLGLEINARDREG-----DTALHDAVRLN 260  
Db 631 yllqhgadnvaqdkggliplnhaasyghdvaaallikynacvnatdkwaftplheaackg 690  
Qy 261 RYKIILKLLHGGADMTKLNLAGKTPDLVQLWQADTRHAL-----EHPE--- 304  
Db 691 rtqlcallahgadtlnkgqgtpdlv---saddvsallitaampsalpsckypqvl 747  
Qy 305 ----PGAENHNGLEGNDSGRETPQVPVA 328  
Db 748 gvrspgatal-----ssgspsspslsa 771

RESULT 14  
AAB27211  
ID AAB27211 standard; Protein; 1166 AA.  
XX  
AC AAB27211;  
XX  
DT 27-FEB-2001 (first entry)  
XX  
DE Human tankyrase II protein sequence SEQ ID NO: 6.  
XX  
KW Human; tankyrase II; telomere length; signal transduction.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 124 /note= "encoded by TTA"  
FT Misc-difference 125 /note= "encoded by TAC"  
XX  
PN WO2000061813-Al.  
XX  
PD 19-OCT-2000.  
XX  
PF 10-APR-2000; 2000WO-US09558.  
XX  
PR 09-APR-1999; 99US-0128577.  
XX  
PR 13-APR-1999; 99US-0129123.  
XX  
PA (GERO-) GERON CORP.  
XX  
PI Morin GB, Funk WD, Piatyszek MA;  
XX  
DR WPI: 2000-679503/66.  
XX  
DR N-PSDB; AAC66825.  
XX  
PT Novel mammalian Tankyrase II polypeptide and the polynucleotide  
PT encoding the polypeptide useful for modulating or maintaining telomere  
PT length, replicative capacity, apoptosis, chromosome packing or gene  
PT expression  
XX

Claim 4; Fig 4; 52pp; English.

PS The present sequence is a version of the human tankyrase II protein  
 XX sequence. The protein is thought to be involved in signal transduction in  
 CC the cell, and to have binding activity for other telomere-associated  
 CC proteins. It is possible that it plays a role in the regulation of  
 CC telomere length, thus affecting the replicative ability of the cell. The  
 CC protein is useful for ribosylating target proteins, for determining  
 CC tankyrase II binding activity in a sample, and for modulating telomere  
 CC length in a cell.  
 XX  
 SQ Sequence 1166 AA;

Query Match 14.6%; Score 247.5; DB 21; Length 1166;  
 Best Local Similarity 23.4%; Pred. No. 5.3e-15;  
 Matches 105; Conservative 49; Mismatches 133; Indels 161; Gaps 13;  
 QY 24 NEKLGRDTRQKLPMDLLVLEDEKHGAQSAALQKVGQERVRKTS---LDLREIIDVG 79  
 DB 408 sekahnd-----vvevvvkhakvnaIdnl-gqtslhraaycghlqtrlllsyg 456  
 QY 80 GIONLIELR-----KKRQKRDALA-ASHEPPPEPEITGPVDEETFLK----- 123  
 DB 457 cdnliisqgftalqmgnenvvqllqegislgnsadrlleaaakagdvctvklctvqs 516  
 QY 124 -----AAVEGKMVIEKFLADGGADTCDFRFTALHRASLEGHMEILE 167  
 DB 517 vncrdiegrqstplhfaagynrvsvveyllqhgadvhakdkggvplhnaacsyghyevae 576  
 QY 168 KLLDNGATVDFQRLDCTAMHWACRGHLEVVVKLLQSHGA-----D 208  
 DB 577 llvkhgavnnvadlwktplheaakgkyeicklllqhgadptkknrdgntpldvkdgd 636  
 QY 209 TNVRDKLL-----SPLHVAVRTGQVEIVE 233  
 DB 637 tdiqdlirgdaalldaakkgclarvkklsppdnvncrdtqgrhstplhlaagynnlvae 596  
 QY 234 HFLSLGLEINARDREG-----DTALHDAVRLN 260  
 DB 697 yllqhgadvnaqdkgglliplhnaasyghvdaallikynacvnatdkwaftplheaakg 756  
 QY 261 RYKIITKLLHGGADMTKNLAGKTPDVLVQWADTRHAL-----EHPE--- 304  
 DB 757 rtqicalllhaagadptlknegqgtpldlv---saddvsalltaampspalpsckypqvin 813  
 QY 305 ----PGAENHGLEPNDSGRETPOVPVA 328  
 DB 814 gvrspgatadal-----ssgspsspslsa 837

RESULT 15  
 AAB66295  
 D AAB66295 standard; Protein; 1166 AA.  
 AC AAB66295;

XX 05-APR-2001 (first entry)  
 XX Human tankyrase2 TANK2-SHORT SEQ ID NO: 135.  
 XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
 KW inflammatory disorder.  
 XX Homo sapiens.  
 OS  
 XX WO200100849-A1.  
 PN  
 XX 04-JAN-2001.  
 PD  
 XX 28-JUN-2000; 2000WO-US17827.  
 PF  
 XX 29-JUN-1999; 99US-0141582.  
 PR

XX (ICOS-) ICOS CORP.

PA Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 PI WPI; 2001-102896/11.  
 DR N-PSDB; AAF63953.  
 XX

XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
 PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
 PT inflammatory and autoimmune disorders  
 XX

XX Claim 3; Page 200-203; 242pp; English.

XX The present invention provides the protein and coding sequence for the  
 CC human tankyrase2 protein. This is found in two different versions,  
 CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
 CC polyADP-ribosylation activity and is involved in the modification of  
 CC TRF1, which is a telomere-specific binding protein. The regulation of  
 CC telomere length, in which TRF1 has a role, is linked to ageing and  
 CC cancer. The sequences are useful in the treatment of cancers and  
 CC inflammatory disorders.  
 XX

SQ Sequence 1166 AA;

Query Match 14.6%; Score 247.5; DB 22; Length 1166;  
 Best Local Similarity 23.4%; Pred. No. 5.3e-15;  
 Matches 105; Conservative 49; Mismatches 133; Indels 161; Gaps 13;

QY 24 NEKLGRDTRQKLPMDLLVLEDEKHGAQSAALQKVGQERVRKTS---LDLREIIDVG 79  
 DB 408 sekahnd-----vvevvvkhakvnaIdnl-gqtslhraaycghlqtrlllsyg 456  
 QY 80 GIONLIELR-----KKRQKRDALA-ASHEPPPEPEITGPVDEETFLK----- 123  
 DB 457 cdnliisqgftalqmgnenvvqllqegislgnsadrlleaaakagdvctvklctvqs 516  
 QY 124 -----AAVEGKMVIEKFLADGGADTCDFRFTALHRASLEGHMEILE 167  
 DB 517 vncrdiegrqstplhfaagynrvsvveyllqhgadvhakdkggvplhnaacsyghyevae 576  
 QY 168 KLLDNGATVDFQRLDCTAMHWACRGHLEVVVKLLQSHGA-----D 208  
 DB 577 llvkhgavnnvadlwktplheaakgkyeicklllqhgadptkknrdgntpldvkdgd 636  
 QY 209 TNVRDKLL-----SPLHVAVRTGQVEIVE 233  
 DB 637 tdiqdlirgdaalldaakkgclarvkklsppdnvncrdtqgrhstplhlaagynnlvae 696  
 QY 234 HFLSLGLEINARDREG-----DTALHDAVRLN 260  
 DB 697 yllqhgadvnaqdkgglliplhnaasyghvdaallikynacvnatdkwaftplheaakg 756  
 QY 261 RYKIITKLLHGGADMTKNLAGKTPDVLVQWADTRHAL-----EHPE--- 304  
 DB 757 rtqicalllhaagadptlknegqgtpldlv---saddvsalltaampspalpsckypqvin 813  
 QY 305 ----PGAENHGLEPNDSGRETPOVPVA 328  
 DB 814 gvrspgatadal-----ssgspsspslsa 837

Search completed: April 17, 2002, 08:10:16  
 Job time: 34 sec

---

THIS PAGE BLANK (USPTO)

**THIS PAGE BLANK (uspto)**



ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C:Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: C86212  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-543 <STO>  
A:Cross-references: GB:AE005172; NID:98439897; PIDN:AAF75083.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 14.6%; Score 247.5; DB 2; Length 543;  
Best Local Similarity 29.1%; Pred. No. 1.4e-09;  
Matches 102; Conservative 48; Mismatches 121; Indels 79; Gaps 15;

QY 21 EENEKLRGTRQKLPK-----DLLVLEDEKHHG-----AQAALOKVKG- 60  
Db 2 EGEDTVAGSSIPKKMMKQLTGKRDITLHSAVRHGNKDRVVEILTTRSELNQLLGK 61  
QY 61 QERVKTSIDLREITDVGGION-----LIELRKKRKKRKKDALAASHEPPPEPEEI 112  
Db 62 QNOSGETALVAAEYGDVEIVKEMINCYDIALVEI-----KARNGFDAFH----- 106  
QY 113 TGPVDEETFLKAAVEGKMKVIEKFLADGGS--ADTCDFQFRRTALHRASLEGHMEILEKLL 170  
Db 107 -----IAKQGDLDVL-KVLAESAHELAMTVLDSNTTALTHTAATQGTVEVNFLL 155  
QY 171 DNGATV-DFODRLDCTAMHWACRGHLEVVK-LLOSHGADTVNRKLLSTPLHVAVRTQ 228  
Db 156 ELGSSLAGIAKNGKTALHSASRNGHVKIKALLASEPAITATRMKKGTQALHMAVKGYN 215  
QY 229 VEIVHEFLSLG-LEINARDREGDTALHDVRLNRYKIIKLLILHG-ADMMTKNLAGKTPT 286  
Db 216 VEVVEELIKADRSINTADTKGNFTALHIAARKGRSQIVKLLANNMTDIAKYNRSGET-- 273  
QY 287 DLVOLWQADTRHAEHP-----PGAENHGLEGNDSGRETPQPV 326  
Db 274 -----ALDTAEKIGNPEVALILQKHGVPASAKTIKPSGNPA-RELKQTV 316

RESULT 3  
T13940  
ankyrin - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
C:Accession: T13940  
Dubreuil, R.K.; Yu, J.  
Proc. Natl. Acad. Sci. U.S.A. 91, 10285-10289, 1994  
A:Title: Ankyrin and beta-spectrin accumulate independently of alpha-spectrin in Drosophila  
A:Reference number: Z17820; MUID:95024098  
A:Accession: T13940  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1549 <DUB>  
A:Cross-references: EMBL:LJ5601; NID:9557083; PID:9557084; PIDN:ANC37208.1  
C:Genetics:  
A:Cross-references: FlyBase:FBgn0011747

Query Match 14.1%; Score 238.5; DB 2; Length 1549;  
Best Local Similarity 25.0%; Pred. No. 2.2e-08;  
Matches 80; Conservative 51; Mismatches 120; Indels 69; Gaps 6;

QY 18 LAQEEENEKLV-----RGDRQKLPMDLLVLEDEKHHGAQAALOKVKGQERVKTSIDLRR 74  
Db 451 LIQHEASADLPTIRGET-----PLHAAANQADIIRILLRSKVDIAIVREGQTPHVASR 506

QY 75 IDVGIGIQNLIELRKKRKKOKKRDALAAASHEPPPEPEITGPVDEETFLKAAVEGKMKVIE 134  
Db 507 LGNINIIMLLQHGAEINAQSNKYSALH-----IAAKEGQENIVQ 547  
QY 135 KFLADGGSADTCDFQFRRTALHRASLEGHMEILEKLLDNGATVDFODRLDCT----- 185  
Db 548 VLEENGAENNAVTKKGFTPLHLACKYGNVVOILLONGASIDFGKNKDVTPLHVATHYN 507  
QY 186 -----AMHWACRGHLEVVKLLQSHGADTVNRKLLSTPLH 221  
Db 608 NPSTIVELLKNGSPNLCARNGQCAIHIACKKNYLSIAMOLLOHGADVNIISKGSFPLH 667  
QY 222 VAVRTGQVEIVHEFLSLGLEINARDREGDTALHDVRLNRYKIIKLLHGDMMTKNLA 281  
Db 668 LAAQGGNVDMVQLLEYGV-ISAANKGLTPLHVAQAQEGHVLVSQILLHGANISERTRN 726  
QY 282 GKTP-----TDLVQLW 292  
Db 727 GYTPLHMAAHYGHLDLVKFF 746

RESULT 4  
D84448  
probable ankyrin [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: D84448  
R.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
cuss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: D84448  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-247 <STO>  
A:Cross-references: GB:AE002093; NID:94335756; PIDN:AAD17433.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g03430  
A:Map position: 2

Query Match 14.0%; Score 237.5; DB 2; Length 247;  
Best Local Similarity 34.1%; Pred. No. 2.6e-09;  
Matches 61; Conservative 32; Mismatches 75; Indels 11; Gaps 2;

QY 117 DEETFLKAAVEGKMKVIEKFLADGGSADTC-----DQFRRTALHRASLEGHMEILEKLLDN 172  
Db 46 DGRSLHVAASFGHSQIVKLLSSDEAKTVINSKDEGWAPLHSAASIGNAELVEVLLTR 105  
QY 173 GATVDFQDRDLCTAMHWACRGHLEVVKLLQSHGADTVNRKLLSTPLHVAVRTGOEIV 232  
Db 106 GADVNAKNGNGRTALHYAASKGRLEIAQLLLTHGAKINITDKVGCTPLHRAASVGKLEVC 165  
QY 233 EHFSLGLEINARDREGDTALHDV-----RLNRYKIIKLLHGDMMTKNLAGKT 284  
Db 166 EFLIEGAEIDATDKMGQTALMHSVICDDKQLKVSMDQVAFLLIRHGADVVDVEDEGYT 224

RESULT 5  
T43458  
hypothetical protein DKFZp434F0621.1 - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
C:Accession: T43458  
R.; Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, December 1999  
A:Reference number: Z22517  
A:Accession: T43458  
A:Status: preliminary  
A:Molecule type: mRNA



A:Residues: 1-1031 <AA>  
A:Cross-references: EMBL:AL133620  
A:Experimental source: adult testis; clone DKFZp434F0621  
C:Genetics:  
A:Note: DKFZp434F0621.1

```
Query Match          13.9%; Score 235; DB 2; Length 1031;
Best Local Similarity 33.5%; Pred. No. 2.3e-08;
Matches 72; Conservative 33; Mismatches 68; Indels 42; Gaps 6;

QY 116 VDEE-----TFKAAVE-GKMKVIEKFLADGGSADTCDFRR-TALHRASLEGHMEILK 169
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 VDQEGANSMTALIVAKGYTQSVKEILRNPNVNLTKDNTALMIASKEGTEIVQDL 256
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      170 LONGATVDFQDLDTAMHWACRGHLEVV-KLQSH-----GAD 208
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 LDQGYVNPIDKSGDTVLIGAVRGGHVEIVRALLOKYADIDTRGDQNKATLVAVKEGNA 316
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 209 TNVRDKLL-----STPLHVAVRTGQVEIVEHFLSLGLSEINARDREGDTALHDA 256
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 317 TNVRDLQCPDTEICTRDGETPLIKATKMRNIEVVELLLDKGAKVSAVDKKGDTPLHTA 376
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 257 VRLNRYKIILLHAGD---MWTNKLAKGTPDLD 288
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 377 IGRSRKLAELLRNPKDGRLLRPNKAGETPYN 411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
B35049
Ankyrin 1, erythrocyte splice form 3 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 10-Jul-1998
C:Accession: B35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: B35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1856 <LAM>
A:Cross-references: GB:M28880
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing
F:2-1856/Product: ankyrin 1, erythrocyte form 3 #status predicted <MAT>
F:2-1513.1676-1856/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
```

```
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match          13.9%; Score 234.5; DB 2; Length 1856;
Best Local Similarity 33.5%; Pred. No. 5.1e-08;
Matches 60; Conservative 29; Mismatches 87; Indels 3; Gaps 2;

QY 108 EPEITGPVDETFKAAVEGKMKVIEKFLADGGSADTCDFRR-TALHRASLEGHMEIL 166
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 330 EDDIT--LDHITPLHVAHCCHRRVAKVLLDKGAKPNSRALNGTPLHIACKKHVRVM 387
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 167 EXLLDNGATVDFQDLDTAMHWACRGHLEVVYKLLQSHGADTNVRDKLLSTPLHVAVRT 226
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 388 ELLKGTASIDAVTESGLTPLHVAFMGHLPIVKNLLQRGASPNVSNVAVETPLHMAARA 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 227 GQVEIVEHFLSLGLEINARDREGDTALHDAVRLNRYKIILLHAGDMMTNLACKTP 285
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 448 GTEVAKYLLQNKAKVNAKDDQTLPHCAARIGHTMVMVKLLLENANPNLATAGTTP 506
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
A35049
Ankyrin 1, erythrocyte splice form 2 - human
N:Alternate names: ankyrin 2.1, erythrocyte; ankyrin-R
N:Contains: ankyrin 2.2, erythrocyte
C:Species: Homo sapiens (man)
C:Date: 27-Jul-1990 #sequence_revision 01-Oct-1992 #text_change 04-Sep-1998
C:Accession: A35049
R:Lambert, S.; Yu, H.; Prchal, J.T.; Lawler, J.; Ruff, P.; Speicher, M.C.;
Proc. Natl. Acad. Sci. U.S.A. 87, 1730-1734, 1990
A:Title: cDNA sequence for human erythrocyte ankyrin.
A:Reference number: A35049; MUID:90175370
A:Accession: A35049
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1880 <LAM>
A:Cross-references: GB:M28880
C:Genetics:
A:Gene: GDB:ANK1; ANK
A:Cross-references: GDB:118737; OMIM:182900
A:Map position: 8p11.2-8p11.2
C:Superfamily: ankyrin; ankyrin repeat homology
C:Keywords: alternative splicing; cytoskeleton
F:2-1880/Product: ankyrin 1, erythrocyte form 2 #status predicted <MAT>
F:2-1513.1676-1880/Product: ankyrin 2.2, erythrocyte #status predicted <MA2>
F:44-76/Domain: ankyrin repeat homology <AN01>
F:77-109/Domain: ankyrin repeat homology <AN02>
F:110-142/Domain: ankyrin repeat homology <AN03>
F:143-171/Domain: ankyrin repeat homology <AN04>
F:172-204/Domain: ankyrin repeat homology <AN05>
F:205-237/Domain: ankyrin repeat homology <AN06>
F:238-270/Domain: ankyrin repeat homology <AN07>
F:271-303/Domain: ankyrin repeat homology <AN08>
F:304-336/Domain: ankyrin repeat homology <AN09>
F:337-369/Domain: ankyrin repeat homology <AN10>
F:370-402/Domain: ankyrin repeat homology <AN11>
F:403-435/Domain: ankyrin repeat homology <AN12>
F:436-468/Domain: ankyrin repeat homology <AN13>
F:469-501/Domain: ankyrin repeat homology <AN14>
F:502-534/Domain: ankyrin repeat homology <AN15>
F:535-567/Domain: ankyrin repeat homology <AN16>
F:568-600/Domain: ankyrin repeat homology <AN17>
F:601-633/Domain: ankyrin repeat homology <AN18>
F:634-666/Domain: ankyrin repeat homology <AN19>
F:667-699/Domain: ankyrin repeat homology <AN20>
F:700-732/Domain: ankyrin repeat homology <AN21>
F:733-765/Domain: ankyrin repeat homology <AN22>
F:766-798/Domain: ankyrin repeat homology <AN23>

Query Match          13.9%; Score 234.5; DB 2; Length 1880;
Best Local Similarity 33.5%; Pred. No. 5.2e-08;
```



A; Residues: 1-1848 <BIR>  
A; Cross-references: EMBL:X59063; NID:g311816; PIDN:CAA48801.1; PID:g311817  
C; Superfamily: ankyrin; ankyrin repeat homology  
C; Keywords: alternative splicing  
F; 48-80/Domain: ankyrin repeat homology <AN01>  
F; 81-113/Domain: ankyrin repeat homology <AN02>  
F; 114-146/Domain: ankyrin repeat homology <AN03>  
F; 147-175/Domain: ankyrin repeat homology <AN04>  
F; 176-208/Domain: ankyrin repeat homology <AN05>  
F; 209-241/Domain: ankyrin repeat homology <AN06>  
F; 242-274/Domain: ankyrin repeat homology <AN07>  
F; 275-307/Domain: ankyrin repeat homology <AN08>  
F; 308-340/Domain: ankyrin repeat homology <AN09>  
F; 341-373/Domain: ankyrin repeat homology <AN10>  
F; 374-406/Domain: ankyrin repeat homology <AN11>  
F; 407-439/Domain: ankyrin repeat homology <AN12>  
F; 440-472/Domain: ankyrin repeat homology <AN13>  
F; 473-505/Domain: ankyrin repeat homology <AN14>  
F; 506-538/Domain: ankyrin repeat homology <AN15>  
F; 539-571/Domain: ankyrin repeat homology <AN16>  
F; 572-604/Domain: ankyrin repeat homology <AN17>  
F; 605-637/Domain: ankyrin repeat homology <AN18>  
F; 638-670/Domain: ankyrin repeat homology <AN19>  
F; 671-703/Domain: ankyrin repeat homology <AN20>  
F; 704-736/Domain: ankyrin repeat homology <AN21>  
F; 737-769/Domain: ankyrin repeat homology <AN22>  
F; 770-802/Domain: ankyrin repeat homology <AN23>

Db	392	ELLTKTASIDAVTESGUTPLHVSFMGHPPLVKNLLQRGASPNVSNVKVETPLHHAARA	451
Qy	227	GOVEIVEHFLSLGLEINARDREGDTALHDVAVLNRYKIIKLLLLHGADMMTNLAGKTP	285
Db	452	GHTWEAKYLLONKAKANAKAKDDOPLHCAARTIGHTGMVKLLLENGASPLNLTATGHTP	510

[illegible]

A: Residues: 1-1862 <RES>  
A/Cross-references: GB:M84756; NID:gl91939; PIDN:AAA37236.1; PID:gl91940  
C: Genetics:  
A: Gene: Ank-1  
C: Superfamily: ankyrin; ankyrin repeat homology  
C: Keywords: alternative splicing  
F: 40-72/Domain: ankyrin repeat homology <AN01>  
F: 73-105/Domain: ankyrin repeat homology <AN02>  
F: 106-138/Domain: ankyrin repeat homology <AN03>  
F: 139-167/Domain: ankyrin repeat homology <AN04>  
F: 168-200/Domain: ankyrin repeat homology <AN05>  
F: 201-233/Domain: ankyrin repeat homology <AN06>  
F: 234-266/Domain: ankyrin repeat homology <AN07>  
F: 267-299/Domain: ankyrin repeat homology <AN08>

107 PEPEITG--PVDEETELKAAVEGKWKVIEKFLADGGSADFCDOFRTALHRASLEGHME 164

C;Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000  
C:Accession: T42713

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.  
J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene  
the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42713

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1943 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710550; PIDN:AAB01606.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

A:Introns: 855/1

C:Function:

A:Description: supposed to play an important role in the polarized distribution of many

A:Note: major kidney ankyrin

C:Superfamily: ankyrin; ankyrin repeat homology

C:Keywords: alternative splicing

Query Match 13.6%; Score 230; DB 2; Length 1943;  
Best Local Similarity 25.9%; Pred. No. 1.1e-07;  
Matches 78; Conservative 41; Mismatches 124; Indels 58; Gaps 7;

QY 22 EENEKLRGDTROKL-PMDLLVLEDEKHHGAQSAALQKVKQGERVVKTSIDLRLREIID--- 77  
Db -272 DRGAKIDAKTRDGLTPL-----HCGARS-----GHEQVVEMLDLSAPILSKTK 315

QY 78 -----VGGIQLNLIELKRRKKQKRDALAASH-----EPP 106

Db 316 NGLSPLHMATQGDHLNCVQLLQHNPVDDVTNDYLTALHVAACHGHHYKAKVLLDKKAS 375

QY 107 PEPEITG--PVDEETFLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHME 164

Db 376 PNAKALNGFTPLH-----IACKKNRIRVMELLLKHGASIQAVTESGLTPIHVAAFMGHVN 430

QY 165 ILEKLLDNGATVDFDRDLCTAMHWACRGHLEVVKLQSHGADTNVRDKLLSTPLHVA 224

Db 431 IVSOLMHHGASPTNVRGETALHMAARSGQAEVRYLVQDGAQVEAKAKDDQTPHISA 490

225 RTGQVEIVEHFLSLGLEINARDREGDTALHDVRLNRYKIKLLLLHGADMMTKNLAKGT 284

Db 491 RLKGADIVQOLLQOGASPNAAATTSVYTPHLAAREGHEDVAAFLLDHGASLSITTTKKGFT 550

QY 285 P 285

Db 551 P 551

RESULT 15

T42716

ankyrin 3, splice form 4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 04-Mar-2000

C:Accession: T42716

R:Peterson, L.L.; John, K.M.; Lu, F.M.; Eicher, E.M.; Higgins, A.; Yialamas, M.; Turtzo, J.

J. Cell Biol. 130, 313-330, 1995

A:Title: Ank3 (epithelial ankyrin), a widely distributed new member of the ankyrin gene

the repeat domain.

A:Reference number: Z22237; MUID:95340633

A:Accession: T42716

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1961 <PET>

A:Cross-references: EMBL:L40632; NID:g710548; PID:g710552; PIDN:AAB01607.1

A:Experimental source: strain C57BL/6J; kidney

C:Genetics:

A:Gene: Ank3

A:Map position: 10

C:Superfamily: ankyrin; ankyrin repeat homology  
C:Keywords: alternative splicing

Query Match 13.6%; Score 230; DB 2; Length 1961;

Best Local Similarity 25.9%; Pred. No. 1.1e-07;

Matches 78; Conservative 41; Mismatches 124; Indels 58; Gaps 7;

QY 22 EENEKLRGDTROKL-PMDLLVLEDEKHHGAQSAALQKVKQGERVVKTSIDLRLREIID--- 77

Db 272 DRGAKIDAKTRDGLTPL-----HCGARS-----GHEQVVEMLDLSAPILSKTK 315

QY 78 -----VGGIQLNLIELKRRKKQKRDALAASH-----EPP 106

Db 316 NGLSPLHMATQGDHLNCVQLLQHNPVDDVTNDYLTALHVAACHGHHYKAKVLLDKKAS 375

QY 107 PEPEITG--PVDEETFLKAAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHME 164

Db 376 PNAKALNGFTPLH-----IACKKNRIRVMELLLKHGASIQAVTESGLTPIHVAAFMGHVN 430

QY 165 ILEKLLDNGATVDFDRDLCTAMHWACRGHLEVVKLQSHGADTNVRDKLLSTPLHVA 224

Db 431 IVSOLMHHGASPTNVRGETALHMAARSGQAEVRYLVQDGAQVEAKAKDDQTPHISA 490

QY 225 RTGQVEIVEHFLSLGLEINARDREGDTALHDVRLNRYKIKLLLLHGADMMTKNLAKGT 284

Db 491 RLKGADIVQOLLQOGASPNAAATTSVYTPHLAAREGHEDVAAFLLDHGASLSITTTKKGFT 550

QY 285 P 285

Db 551 P 551

Search completed: April 17, 2002, 08:11:00

Job time: 78 sec

THIS PAGE BLANK (USPTO)

result No.	Score	Query			DB	ID	Description
		Match	Length				
1	284.5	16.8	1059	1	Y379_HUMAN	O15084	homo sapien
2	245	14.5	692	1	Y9294_HUMAN	Q9Y294	homo sapien
3	234.5	13.9	1880	1	ANK1_HUMAN	P16157	homo sapien
4	234.5	13.9	3924	1	ANK2_HUMAN	Q01484	homo sapien
5	233.5	13.8	1862	1	ANK2_MOUSE	Q02357	mus musculus
6	228.5	13.5	318	1	IKBA_CHICK	Q91974	gallus gall
7	224.5	13.3	768	1	YKB23_HUMAN	Q9ulj7	homo sapien
8	224	13.2	679	1	RN5A_MOUSE	Q05921	mus musculus
9	217	12.8	741	1	RN5A_HUMAN	Q05823	homo sapien
10	215	12.7	323	1	ANKH_CHRVI	Q06527	chromatium
11	212	12.5	656	1	FEM1_CAEEL	P17221	caenorhabdi
12	208.5	12.3	347	1	GABC_MOUSE	Q00421	mus musculus
13	208.5	12.3	382	1	GABB_MOUSE	Q00420	mus musculus
14	207.5	12.3	347	1	GABC_HUMAN	Q06545	homo sapien
15	207.5	12.3	383	1	GABK_HUMAN	Q06547	homo sapien
16	203.5	12.0	1431	1	DAPK_HUMAN	P53355	homo sapien
17	202	11.9	832	1	ANR3_HUMAN	P57078	homo sapien
18	200.5	11.8	1401	1	LATA_LATMA	P23631	latrodectus
19	198	11.7	314	1	IKBA_RAT	Q63644	rattus norv
20	196	11.6	314	1	IKBA_PIG	Q08353	sus scrofa
21	189.5	11.2	439	1	AKR_ARATH	Q05753	arabidopsis
22	187.5	11.1	317	1	IKBA_HUMAN	P25963	homo sapien
23	187	11.0	765	1	BAR1_MOUSE	O70445	mus musculus
24	185	10.9	168	1	CDNG_HUMAN	P42773	homo sapien
25	181	10.7	768	1	BAR1_RAT	Q9qzh2	rattus norv
26	180.5	10.7	1083	1	YIL2_YEAST	P04480	saccharomyc
27	179.5	10.6	2703	1	NOTC_DROME	.P07207	drosophila
28	178.5	10.5	764	1	AKR1_YEAST	P39010	saccharomyc
29	178.5	10.5	971	1	KBFL_MOUSE	P25799	mus musculus
30	176.5	10.4	238	1	IKBA_MOUSE	Q921e3	mus musculus
31	176.5	10.4	414	1	GABD_MOUSE	P81069	mus musculus
32	175.5	10.4	522	1	KBFL_RAT	Q03369	rattus norv
33	174	10.3	984	1	KBFL_CHICK	Q04861	gallus gall

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AB023174; BAA76801.1; -  
CC EMBL; AL117504; CAB5968.1; -  
CC EMBL; AL096678; CAB86658.1; -  
CC HSP; Q00421; IAWC.  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 7.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SM00248; ANK; 6.  
DR PROSITE; PS00088; ANK\_REPEAT; 6.  
DR PROSITE; PS0297; ANK\_REPEAT\_REGION; 1.  
KW Hypothetical protein; Repeat; ANK repeat; Alternative splicing.  
FT REPEAT 9 38  
FT REPEAT 41 70  
FT REPEAT 74 103  
FT REPEAT 107 136  
FT REPEAT 140 169  
FT REPEAT 173 202  
FT REPEAT 206 235  
FT REPEAT 239 268  
FT VARSPLIC 264  
FT CONFLICT 87 87  
FT SEQUENCE 692 AA; 75722 MW; D32115487A80839F CRC64;  
Query Match 14.5%; Score 245; DB 1; Length 692;  
Best Local Similarity 32.7%; Pred. No. 1.1e-09;  
Matches 56; Conservative 36; Mismatches 79; Indels 0; Gaps 0;

QY 117 DEETFLKAAVEGKMKVIEKFLADGSGADTCDFRRRTALHRASLEGHMEILEKLLDNGATV 176  
Db 75 DQTAHRAVTVGVNTEIIAALHESCALDRQDKDGTALHEASWHGFSQSAKLLVKGAGNY 134  
QY 177 DFQRLDCTAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFL 236  
Db 135 LAKNKAGNTALHLACONSHSQSTRVLLLAGSRADLKNAGDTCLHVAARYNHLISIRLL 194  
QY 237 SLGLEINARDREGDTALHDVAVLRNRYKIILKLLHGDMMTKNLAKGTPD 287  
Db 195 TAFCSVHEKNGAGDTALHVAALNKKVAKILLEGADTTIVNAGOTPLE 245

RESULT 3  
ID ANK1\_HUMAN STANDARD; PRT; 1880 AA.  
AC P16157;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN) (ANKYRIN R) (ANKYRINS 2.1 AND 2.2).  
GN ANK1 OR ANK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND VARIANTS.  
RC TISSUE=Hematopoietic;  
RX MEDLINE=90158830; PubMed=2137557;  
RA Lux S.E., John K.M., Bennett V.;  
RT "Analysis of cDNA for human erythrocyte ankyrin indicates a repeated  
RT structure with homology to tissue-differentiation and cell-cycle  
RL Nature 344:36-42(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90175370; PubMed=1689849;  
RA Lambert S., Yu H., Prchal J.T., Lawler J., Ruff P., Speicher D.,  
RA Cheung M.C., Kan Y.W., Palek J.;

FT REPEAT 510 540 ANK 15.  
FT REPEAT 555 584 ANK 16.  
FT REPEAT 588 617 ANK 17.  
FT REPEAT 622 651 ANK 18.  
FT REPEAT 658 687 ANK 19.  
FT REPEAT 691 720 ANK 20.  
FT REPEAT 724 753 ANK 21.  
FT REPEAT 761 790 ANK 22.  
FT REPEAT 793 823 ANK 23.  
FT REPEAT 828 857 ANK 24.  
FT REPEAT 861 891 ANK 25.  
FT REPEAT 895 924 ANK 26.  
FT REPEAT 931 960 ANK 27.  
SQ SEQUENCE 1059 AA; 113465 MW; C1F55E6CFE494770 CRC64;

Query Match 16.8%; Score 284.5; DB 1; Length 1059;  
Best Local Similarity 37.3%; Pred. No. 3.5e-12;  
Matches 66; Conservative 32; Mismatches 76; Indels 3; Gaps 1;  
QY 124 AAVEGKMKVIEKFLADGSGADTCDFRRRTALHRASLEGHMEILEKLLDNGATVDFQDRID 183  
Db 120 AAANKAVKCAEALVPLLSNVNVDGRGRTALHHAFAFGHGMVKKLLSRGANINAFDKD 179  
QY 184 CTAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIVEHFLSLGLEIN 243  
Db 180 RRAIHWAAVMGHTEVVKLLVSHGAEVTCCKDKSYTPLHAAASGSMISVVVKYLLDGLGVDNM 239  
QY 244 ARDEGDTALHDVAVLRNRYKIILKLLHGDMMTKNLAKGTPDVLQVQLWQADTRHAL 300  
Db 240 EPNAYGNTPLHVACVNGODVVVVELIDCGAIVNQKNEKGFTP---LHFAAASHTGAL 293

RESULT 2  
ID Y957\_HUMAN STANDARD; PRT; 692 AA.  
AC Q912G4; Q9NU24; Q9UFQ9;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN KIAA0957.  
GN KIAA0957.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Brain;  
RX MEDLINE=99246063; PubMed=10231032;  
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,  
MIYAJIMA N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RL for large proteins in vitro.";  
RL DNA Res. 6:63-70(1999).  
RN [2]  
RP SEQUENCE OF 87-692 FROM N.A. (ISOFORM 2).  
RC TISSUE=Testis;  
RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 140-692 FROM N.A. (ISOFORM 1).  
RA Tracey A.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; MAY BE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: CONTAINS 8 ANK REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way





QY	124	A A V E G K M V I E K F L A D G S A D T C D F R R T A L H R A S L E G H M E I L K L O N G A T V D F Q D R L D	183
		: :     :	:
D b	636	A A K K N O M Q L A S T L L N V G A E T N I V T K G Y P T P L H A S Q B G H T D M V I L L D K G A N I H M S T K S G	695
		: :     :	:
QY	184	C T A M H W A C R G G H L E V V K L L Q S G A D T N V R D K L L S T P L H V A V T G O V E I V E H P L S L G L E I N	243
		: :     :	:
D b	696	L I S L H L A A Q E D K V N V A D I L T K H G A D Q A H T K L G Y T P L I V A C H Y G N V K M V N F L L K O G A N V N	755
		: :     :	:
QY	244	A R D R E G D T A L H D A V R U N R Y K I I K L L L H G A D M M T K N A G K T P D L V Q - - - L W Q A D T - - -	296
		: :     :	:
D b	756	A K T K N G Y T P L H O A O O G C H Y T H I I N V L L O H G A K P N A T T A N G N T A L A K R U G I T S V D V T L K V	815
		: :     :	:

QY 297 -----RHLEHPEPGAEHNGLEPNDSGRET 322  
DB 816 VTVEVTTTTTITKHLNVPETWTE--VLDVSDDEGDDT 853

RESULT 5  
ANK1\_MOUSE STANDARD; PRT; 1862 AA.  
ID ANK1\_MOUSE  
AC Q02357;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ANKYRIN 1 (ERYTHROCYTE ANKYRIN).  
GN ANK1 OR ANK-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.  
RP TISSUE-Erythrocyte;  
RC MEDLINE=92345717; PubMed=1386265;  
RA White R.A., Birkenmeier C.S., Peters L.L., Barker J.E., Lux S.E.;  
RT "Murine erythrocyte ankyrin cDNA: highly conserved regions of the  
regulatory domain";  
RL Mamm. Genome 3:281-285(1992).  
CC -1- FUNCTION: ATTACH INTEGRAL MEMBRANE PROTEINS TO CYTOSKELETAL  
ELEMENTS; BIND TO THE ERYTHROCYTE MEMBRANE PROTEIN BAND 4-2. TO  
NA-K ATPASE, TO THE LYMPHOCYTE MEMBRANE PROTEIN GP85, AND TO THE  
CYTOSKELETAL PROTEINS FODRIN, TUBULIN, VIMENTIN AND DESMIN.  
CC ERYTHROCYTE ANKYRINS ALSO LINK SPECTRIN (BETA CHAIN) TO THE  
CYTOSKELETAL DOMAIN OF THE ERYTHROCYTES ANION EXCHANGE PROTEIN;  
CC THEY RETAIN MOST OR ALL OF THESE BINDING FUNCTIONS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC SURFACE OF ERYTHROCYTIC  
PLASMA MEMBRANE.  
CC -1- PTM: REGULATED BY PHOSPHORYLATION (BY SIMILARITY).  
CC -1- PTM: ACYLATED BY PALMITIC ACID GROUP(S) (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 23 ANK REPEATS.  
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.  
-----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; M84756; AAA37236.1; -  
DR HSSP; Q00420; 1ANC.  
DR MGD; MGI:88024; Ank1.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00023; ank; 24.  
DR Pfam; PF00531; death; 1.  
DR Pfam; PF00791; ZU5; 1.  
DR SMART; SM00248; ANK; 22.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50088; ANK\_REPEAT; 20.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
KW Cytoskeleton; Repeat; ANK repeat; Phosphorylation; Lipoprotein.  
FT DOMAIN 1 827  
FT 89 KDA DOMAIN (ANION EXCHANGE PROTEIN  
FT BINDING DOMAIN).  
FT DOMAIN 828 1386  
FT 62 KDA DOMAIN (SPECTRIN BINDING  
FT DOMAIN).  
FT DOMAIN 1387 1862  
FT 55 KDA REGULATORY DOMAIN (REGULATES  
FT THE BINDING OF ANKYRIN TO SPECTRIN  
FT AND THE BAND 3 PROTEIN).  
FT REPEAT 40 69  
FT REPEAT 73 102

FT REPEAT 106 135 ANK 3.  
FT REPEAT 139 168 ANK 4.  
FT REPEAT 170 197 ANK 5.  
FT REPEAT 201 230 ANK 6.  
FT REPEAT 234 263 ANK 7.  
FT REPEAT 267 296 ANK 8.  
FT REPEAT 300 329 ANK 9.  
FT REPEAT 333 362 ANK 10.  
FT REPEAT 366 395 ANK 11.  
FT REPEAT 399 428 ANK 12.  
FT REPEAT 432 461 ANK 13.  
FT REPEAT 465 494 ANK 14.  
FT REPEAT 498 527 ANK 15.  
FT REPEAT 531 560 ANK 16.  
FT REPEAT 564 593 ANK 17.  
FT REPEAT 597 626 ANK 18.  
FT REPEAT 630 659 ANK 19.  
FT REPEAT 663 692 ANK 20.  
FT REPEAT 696 725 ANK 21.  
FT REPEAT 729 758 ANK 22.  
FT REPEAT 762 791 ANK 23.  
FT DOMAIN 1399 1483 DEATH.  
SQ SEQUENCE 1862 AA; 204242 MW; AE6B85B5B29001E5 CRC64;  
  
Query Match 13.8%; Score 233.5; DB 1; Length 1862;  
Best Local Similarity 34.1%; Pred. No. 2.3e-08;  
Matches 61; Conservative 27; Mismatches 88; Indels 3; Gaps 2;  
  
QY 108 EPEETGPVDETFKAAVEGKMKVIEKFLADGSGADTCDFRR-TALHRSASLEHMEIL 166  
DB 326 EIDIT--LDHLTPLHVAACHGHRVAKVLLDGKAKPNRSLNGFTPLHIACKKNHIRM 383  
QY 167 EKLDNGATVDFQDRLDCTAMHWACRGHLEVKLLQSHGADTNVRDKLLSTPLHVAVRT 226  
DB 384 ELLLTGASIDAVTESGLTPLHVASFMGHLPIVKNLLQRGASPNVSVKVTPLHMAARA 443  
QY 227 GQVEIVEHFLSGLEINARDREGDTALHDVRLNRYKIKILLHGGADMVTKNLAKGTP 285  
DB 444 GHTEVAKYLLONKAKANAKAKDDQDTPLHCAARIGHTGMVKLLLENGASPNLATTAGTTP 502

RESULT 6  
IKBA\_CHICK STANDARD; PRT; 318 AA.  
ID IKBA\_CHICK  
AC Q91974;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE NF-KAPPAB INHIBITOR ALPHA (I-KAPPA-B-ALPHA) (IKAPPABALPHA) (IKB-ALPHA)  
(REL-ASSOCIATED PROTEIN PP40).  
GN NFkBIA OR IKBA.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Embryonic fibroblast;  
RX MEDLINE=91368196; PubMed=1891714;  
RA Davis N., Ghosh S., Simmons D.L., Tempst P., Liou H.-C., Baltimore D.,  
RA Bose H.R. Jr.;  
RT "Rel-associated pp40: an inhibitor of the rel family of transcription  
factors";  
RL Science 253:1268-1271(1991).  
CC -1- FUNCTION: INHIBITOR OF NF-KAPPA-B THAT TIGHTLY REGULATES NF-KAPPA-  
B ACTIVATION BY COMPLEXING AND TRAPPING IT IN THE CYTOPLASM. MAY  
BE INVOLVED IN REGULATION OF TRANSCRIPTIONAL RESPONSES TO NF-  
KAPPA-B, INCLUDING CELL ADHESION, IMMUNE AND PROINFLAMMATORY  
RESPONSES, APOPTOSIS, DIFFERENTIATION AND GROWTH. CONTROLLED BY  
SEQUENTIAL SERINE-PHOSPHORYLATION, UBIQUITINATION AND DEGRADATION.  
CC TYROSINE-PHOSPHORYLATION COULD ONLY LEAD TO DISSOCIATION FROM NF-  
CC



Db 352 RSALQSAANQGHVKKVQLLIEHGAUVVHTCNOGATACALTAQEGHIDVVQVLLHGAADPN 411  
QY 244 ARDEGDTALHDVRLNRKIIKLLHLCADMMTKNLAGKTPDVLQVQWADTRALE 301  
Db 412 HADQFQRTAMRVAAKNGHSQIIKLEKYGA-----SSLNGCCSPV-----HTME 456

RESULT 8  
RNSA\_MOUSE  
ID RNSA\_MOUSE STANDARD; PRT; 679 AA.  
AC Q05921;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 2-5A-DEPENDENT RIBONUCLEASE (EC 3.1.26.-) (2-5A-DEPENDENT RNASE)  
DE (RIBONUCLEASE L) (RNASE L) (RIBONUCLEASE 4) (FRAGMENT).  
GN RNASL OR RNS4.  
Mus musculus (Mouse).  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=93201598; PubMed=7680958;  
RX Zhou A., Hassel B.A., Silverman R.H.;  
RA "Expression cloning of 2-5A-dependent RNase: a uniquely regulated  
RT mediator of interferon action."  
RL Cell 72:753-765(1993).  
CC -1- FUNCTION: MEDIATES THE EFFECTS OF INTERFERON AGAINST  
CC PICORNAVIRUSES. ACTIVE WHEN BOUND TO 2-5A (5'-PHOSPHORYLATED  
CC 2',5'-LINKED OLIGONUCLEOTIDES). CAPABLE OF CLEAVING POLY(RU) AND,  
CC TO A LESSER EXTENT, POLY(RA), TO SETS OF DISCRETE PRODUCTS RANGING  
CC FROM BETWEEN 4 AND 22 NUCLEOTIDES IN LENGTH.  
CC -1- CATALYTIC ACTIVITY: CLEAVES 3' OF UPNP DIMERS, WITH PREFERENCE  
CC FOR UU AND UA SEQUENCES.  
CC -1- COFACTOR: OPTIMAL RNA CLEAVAGE RATES REQUIRES THE PRESENCE OF  
CC EITHER MANGANESE OR MAGNESIUM AND ATP.  
CC -1- SUBUNIT: MONOMER.  
CC -1- INDUCTION: BY INTERFERON.  
CC -1- SIMILARITY: CONTAINS 9 ANK REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L10382; AAA37117.1;  
DR HSP; Q00420; 1AWC.  
DR MGD; MGI:1098272; Rnasel.  
DR InterPro; IPR002110; ANK.  
DR InterPro; IPR007019; Euk\_pkinase.  
DR Pfam; PF00023; ank; 8.  
DR Pfam; PF00069; pkinase; 1.  
DR SMART; SM00248; ANK; 7.  
DR PROSITE; PS50088; ANK\_REPEAT; 7.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Hydrolyase; Nuclease; RNA-binding; Endonuclease; Zinc-finger;  
KW Repeat; ANK repeat.  
KW REPEAT 24 53  
FT REPEAT 58 87 ANK 1.  
FT REPEAT 91 120 ANK 2.  
FT REPEAT 124 153 ANK 3.  
FT REPEAT 167 197 ANK 4.  
FT REPEAT 201 234 ANK 5.  
FT REPEAT 238 268 ANK 6.  
FT REPEAT 272 301 ANK 7.  
FT REPEAT 303 328 ANK 8.  
FT REPEAT 364 584 ANK 9.  
FT DOMAIN 229 242  
FT 2-5A BINDING (P-LOOP) 1.

FT DOMAIN 253 275 2-5A BINDING (P-LOOP) 2.  
FT NON\_FING 401 436 C6-TYPE (POTENTIAL).  
FT NON\_TER 679 679  
SQ SEQUENCE 679 AA; 76800 MW; 507DB36B962EC1F4D CRC64;  
Query Match 13.28; Score 224; DB 1; Length 679;  
Best Local Similarity 28.18; Pred. No. 2.9e-08;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;  
QY 41 VLEDEKHHGAOSAALQVKGOERVKRTSLDLRRLREITVDVGGIQLIE-----LRKKRKQ 93  
Db 22 VVEDD-----SSLIKAVQKGVVRV-----QQLLEKAGADANACEDTWGTPPLHNAVQA 69  
QY 94 KKRDA--LAASHEPPEPEITGPVDETFLLKAAVEGKMKVIEKFLADGGSGADTCQ--- 148  
Db 70 GRVDIVNLLSHGADPHRRKKGAT---PFTAGIQGVKLLLEILLSCGADVNECDNGF 126  
QY 149 -----FR-----TALHRASLEGHMEILEK 168  
Db 127 TAFMEAERGNAAELRFLFAKGVNLRRTQTKRRRLKQGGATALMSAAEKGHLEVLRI 186  
QY 169 LL-DNGATVDFODR-----LDCAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLRV 222  
Db 187 LLNDKAEVDARDNMGRNALIRTLNWDENVE-EITSLIQHGADVNVGRGKTPLIA 245  
QY 223 AVRTGOVEIVEHFLSL-GLEINARREGDTALHDVRLNRKIIKLLHLCADMMTKNLA 281  
Db 246 AVERKHTGLVQMLLSREGINIDARDNEGTALLIAVDKQLKEIVQLLEKADKC----- 300  
QY 282 GKTPDVLQVQWADTRAH 298  
Db 301 ----DDLW--WIARNH 311  
RESULT 9  
RNSA\_HUMAN  
ID RNSA\_HUMAN STANDARD; PRT; 741 AA.  
AC Q05823;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 2-5A-DEPENDENT RIBONUCLEASE (EC 3.1.26.-) (2-5A-DEPENDENT RNASE)  
DE (RIBONUCLEASE L) (RNASE L) (RIBONUCLEASE 4).  
GN RNASL OR RNS4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=93201598; PubMed=7680958;  
RA Zhou A., Hassel B.A., Silverman R.H.;  
RT "Expression cloning of 2-5A-dependent RNase: a uniquely regulated  
RL mediator of interferon action."  
RL Cell 72:753-765(1993).  
RN [2]  
RN CHARACTERIZATION.  
RX MEDLINE=9425737; PubMed=7514601;  
RA Dong B., Xu L., Zhou A., Hassel B.A., Lee X., Torrence P.F.,  
RA Silverman R.H.;  
RT "Intrinsic molecular activities of the interferon-induced 2-5A-  
RT dependent RNase."  
RL J. Biol. Chem. 269:14153-14158(1994).  
CC -1- FUNCTION: MEDIATES THE EFFECTS OF INTERFERON AGAINST  
CC PICORNAVIRUSES. ACTIVE WHEN BOUND TO 2-5A (5'-PHOSPHORYLATED  
CC 2',5'-LINKED OLIGONUCLEOTIDES). CAPABLE OF CLEAVING POLY(RU) AND,  
CC TO A LESSER EXTENT, POLY(RA), TO SETS OF DISCRETE PRODUCTS RANGING  
CC FROM BETWEEN 4 AND 22 NUCLEOTIDES IN LENGTH.  
CC -1- CATALYTIC ACTIVITY: CLEAVES 3' OF UPNP DIMERS, WITH PREFERENCE  
CC FOR UU AND UA SEQUENCES.  
CC -1- COFACTOR: OPTIMAL RNA CLEAVAGE RATES REQUIRES THE PRESENCE OF



```
RESULT 11
ID FEM1_CAEEL STANDARD; PRT; 656 AA.
AC P17221.
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SEX-DETERMINING PROTEIN FEM-1.
GN FEM-1 OR ISX-1 OR F35D6.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=90199879; PubMed=2317869;
Spence A.M., Coulson A., Hodgkin J.;
*The product of fem-1, a nematode sex-determining gene, contains a
found in cell cycle control proteins and receptors for
cell-cell interactions.*;
RL Cell 60:981-990(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SEX-DETERMINING PROTEIN. IT IS ESSENTIAL FOR THE
CC ADOPTION OF THE MALE SEXUAL FATE IN ALL TISSUES IN C.ELEGANS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -!- SIMILARITY: CONTAINS 7 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; J03172; AAA28055.1; -
DR EMBL; U52001; AAA96093.1; -
DR PIR; A34793; A34793.
DR HSP; Q00420; LAWC.
DR WormPep; F35D6.1; CE07175.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 8.
DR SMART; SM00248; ANK; 5.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 2.
DR Developmental protein; Phosphorylation; ANK repeat; Repeat.
FT REPEAT 47 76 ANK 1.
FT REPEAT 88 118 ANK 2.
FT REPEAT 122 151 ANK 3.
FT REPEAT 155 184 ANK 4.
FT REPEAT 188 217 ANK 5.
FT REPEAT 220 250 ANK 6.
FT REPEAT 557 588 ANK 7.
SQ SEQUENCE 656 AA; 74286 MW; 4F5FC3FA3EA950D3 CRC64;

Query Match 12.5%; Score 212; DB 1; Length 656;
Best Local Similarity 32.7%; Pred. No. 1.9e-07;
Matches 56; Conservative 30; Mismatches 75; Indels 10; Gaps 3;

QY 124 AAVEGKMKVIEKFLADGSGA-----DRCDFRRTALHRASLEGHMELEKLLDN-GA 174
|| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 55 AARGHANVVEYLLEIGADPSVRGWEFDNENIQGTPLWAAASAGHIEVTKLLIEKANA 114
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 175 TVDFQRLDCTAMHWACRGHLEVVKLLQSHGADTNVRKLLSTPLHVAVRTGQVEIVEH 234
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 115 DVNOATWTRSTPLRGACYDGHLDIVKYLEKGADPHIPNRHGHTCLMIASRYRNKVGIVEE 174
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 235 FLSLGLEINARDREGDTALHDAVRLNRYKIIKLLHLLHGDMMTKNLACKTP 285
| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 175 LLKTGIDVKKTERGNTALHDAESGNVEVVKILLKHGSLVM-KDIOGVDP 224

RESULT 12
ID GABC_MOUSE STANDARD; PRT; 347 AA.
AC Q00421;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GA BINDING PROTEIN BETA-2 CHAIN (GABP-BETA-2 SUBUNIT) (GABBP2).
GN GABPB1 OR GABPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91343912; PubMed=1876836;
Lamarco K., Thompson C.C., Byers B.P., Walton E.M., McKnight S.L.;
*Identification of Ets- and notch-related subunits in GA binding
protein.*;
RL Science 253:789-792(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 5-157.
RX MEDLINE=98128030; PubMed=9461436;
Batchelor A.H., Piper D.E., de la Brousse F.C., McKnight S.L.,
Wolberger C.;
*The structure of GABPalpha/beta: an ETS domain-ankyrin repeat
heterodimer bound to DNA.*;
RL Science 279:1037-1041(1998).
CC -!- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE
CC RICH REPEATS (GA REPEATS).
CC -!- SUBUNIT: HETEROETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -!- ALTERNATIVE PRODUCTS: THE BETA-1 AND BETA-2 SUBUNITS ARE
CC PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY
CC DIFFER IN THEIR C-TERMINAL EXTREMITY.
CC -!- SIMILARITY: CONTAINS 5 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M74517; AAA53032.1; -
DR PIR; C40858; C40858.
DR PDB; LAWC; 18-MAR-98.
DR TRANSFAC; T00298; -.
DR TRANSFAC; T01404; -.
DR MGD; MGI:95611; Gabpbl.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 3.
DR SMART; SM00248; ANK; 3.
DR PROSITE; PS50088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR Transcription regulation; Nuclear protein; Alternative splicing;
KW ANK repeat; Repeat; 3d-structure.
FT REPEAT 5 34 ANK 1.
FT REPEAT 37 66 ANK 2.
FT REPEAT 70 99 ANK 3.
FT REPEAT 103 132 ANK 4.
FT REPEAT 136 166 ANK 5.
FT CONFLICT 233 233 V -> VV (IN REF. 1; AAA53032).
SQ SEQUENCE 347 AA; 36867 MW; 01D6AB96B8B83D05 CRC64;

Query Match 12.3%; Score 208.5; DB 1; Length 347;
Best Local Similarity 27.8%; Pred. No. 1.4e-07;
```



Matches 64; Conservative 34; Mismatches 101; Indels 31; Gaps 5;

QY 75 IIDVGGIQLNLELRKKRQKRDALAAASHEPPEPPEITGPVDEETFLKAAVEGKMKVIE 134  
 Db 3 LVDLG--KKLEAARAGQDDEVIRILMANGAPFTDWTGTSPLH-----LAAQYGHFSTTE 55

QY 135 KFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFODRLDCTAMHWACRGG 194  
 Db 56 VLLRAGVSRDARTKVDRTPLHMAASEGHANIVEVLLKHGADVNAKMDLKMALHWATEHN 115

QY 195 HLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIHEHF--LSLGLLEINARDREGDTAL 253  
 Db 116 HQEVVELLIKYGADVHTQSKFCKTAFDISIDNGEDLAEILQIAMQNIINTNPESPT-- 173

QY 254 HDVRLNRYKIILKLLHGDMM-----TKNLAGTKPTDVLQWQADT 296  
 Db 174 -----VTIHAATPQFIIGPGVNVLTDETVSAVQFGNSST 209

RESULT 13

ID GABB\_MOUSE STANDARD; PRT; 382 AA.

AC Q00420;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE GA BINDING PROTEIN BETA-1 CHAIN (GABP-BETA-1 SUBUNIT) (GABPB1).

GN GABPB1 OR GABPB.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91343912; PubMed=1876836;

RA Lamarco K., Thompson C.C., Byers B.P., Walton E.M., McKnight S.L.;  
 RT "Identification of Ets- and notch-related subunits in GA binding  
 protein.";  
 RL Science 253:789-792(1991).

RN [2]

RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 5-157.

RX MEDLINE=98128030; PubMed=9461436;

RA Batchelor A.H., Piper D.E., de la Brousse F.C., McKnight S.L.,  
 RA Wolberger C.;  
 RT "The structure of GABPalpha/beta: an ETS domain-ankyrin repeat  
 heterodimer bound to DNA.";  
 RL Science 279:1037-1041(1998).

CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE  
 CC RICH REPEATS (GA REPEATS).

CC -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.

CC -1- ALTERNATIVE PRODUCTS: THE BETA-1 AND BETA-2 SUBUNITS ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY  
 CC DIFFER IN THEIR C-TERMINAL EXTREMITY.

CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M74516; AAA53031.1; -  
 CC PIR; B40858; B40858.  
 CC PDB; 1AUC; 18-MAR-98.  
 CC TRANSFAC; T00298; -  
 CC TRANSFAC; T01403; -  
 CC MGD; MGI:95611; Gabpb1.  
 CC InterPro; IPR002110; ANK.  
 CC Pfam; PF00023; ank; 3.  
 CC SMART; SM00248; ANK; 3.  
 CC PROSITE; PS50088; ANK\_REPEAT; 3.

DR PROSITE; PS50297; ANK\_REPEAT; 1.  
 KW transcription regulation; Nuclear protein; Alternative splicing;  
 FT REPEAT 5 34 ANK 1.  
 FT REPEAT 37 66 ANK 2.  
 FT REPEAT 70 99 ANK 3.  
 FT REPEAT 103 132 ANK 4.  
 FT REPEAT 136 166 ANK 5.  
 SQ SEQUENCE 382 AA; 41258 MW; 0A5FE3F71D4227AF CRC64;

Query Match 12.3%; Score 208.5; DB 1; Length 382;  
 Best Local Similarity 27.8%; Pred. No. 1.6e-07;  
 Matches 64; Conservative 34; Mismatches 101; Indels 31; Gaps 5;

QY 75 IIDVGGIQLNLELRKKRQKRDALAAASHEPPEPPEITGPVDEETFLKAAVEGKMKVIE 134  
 Db 3 LVDLG--KKLEAARAGQDDEVIRILMANGAPFTDWTGTSPLH-----LAAQYGHFSTTE 55

QY 135 KFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFODRLDCTAMHWACRGG 194  
 Db 56 VLLRAGVSRDARTKVDRTPLHMAASEGHANIVEVLLKHGADVNAKMDLKMALHWATEHN 115

QY 195 HLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIHEHF--LSLGLLEINARDREGDTAL 253  
 Db 116 HQEVVELLIKYGADVHTQSKFCKTAFDISIDNGEDLAEILQIAMQNIINTNPESPT-- 173

QY 254 HDVRLNRYKIILKLLHGDMM-----TKNLAGTKPTDVLQWQADT 296  
 Db 174 -----VTIHAATPQFIIGPGVNVLTDETVSAVQFGNSST 209

RESULT 14

ID GABC\_HUMAN STANDARD; PRT; 347 AA.

AC Q06545;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE GA BINDING PROTEIN BETA-2 CHAIN (GABP-BETA-2 SUBUNIT) (TRANSCRIPTION  
 DE FACTOR E4TF1-47) (GABPB2).

GN GABPB1 OR GABPB OR E4TF1B.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93180783; PubMed=8441384;

RA Watanabe H., Sawada J.-I., Yano K.-I., Yamaguchi K., Goto M.,  
 RA Handa H.;  
 RT "cDNA cloning of transcription factor E4TF1 subunits with Ets and  
 RT notch motifs";  
 RL Mol. Cell. Biol. 13:1395-1391(1993).

CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE  
 CC RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE  
 CC ADENOVIRUS E4 GENE.

CC -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.

CC -1- ALTERNATIVE PRODUCTS: THE BETA-1 AND BETA-2 SUBUNITS ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY  
 CC DIFFER IN THEIR C-TERMINAL EXTREMITY.

CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; D13316; BAA02573.1; -  
 CC HSSP; Q00420; IAWC.



DR TRANSFAC; T01392; -  
DR MIN; 600610; -  
DR InterPro: IPR002110; ANK.  
DR Pfam; PF00023; ank; 3.  
DR SMART; SM00248; ANK; 3.  
DR PROSITE; PS50088; ANK\_REPEAT; 3.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
KW Transcription regulation; Nuclear protein; Alternative splicing;  
KW ANK repeat; Repeat. 34 ANK 1.  
FT REPEAT 37 66 ANK 2.  
FT REPEAT 70 99 ANK 3.  
FT REPEAT 103 132 ANK 4.  
FT REPEAT 136 166 ANK 5.  
SQ SEQUENCE 347 AA; 36850 MW; 902C02A3F6ABBA2E CRC64;

Query Match 12.3%; Score 207.5; DB 1; Length 347;  
Best Local Similarity 27.8%; Pred. No. 1.7e-07;  
Matches 64; Conservative 34; Mismatches 101; Indels 31; Gaps 5;

Qy 75 IIDVGGIONLIELKRRKQKRDALAAASHEPPEPEETGPVDETFLLKAAVEGKMKVIE 134  
Db 3 LVDLG--KKLLEAARAGQDDEVIRILMANGAPFTTDWLGTSPH-----LAAQYGHYSTTE 55  
Qy 135 KFLADGGSADTCQDPRRTALHRASLEGHMEILEKLLDNGATVDFODRLDCTAMHWACRG 194  
Db 56 VLLRAGVSRDARTKVDRTPLHMAASEGHASIVEVLKKGADVNKADMLKMTALHWATEHN 115  
Qy 195 HLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIYEHF-LSGLGLEINARDREGDTAL 253  
Db 116 HOEVVELLIKYGADVHTQSKFCKTAFDISIDNGNEDLAEILQIAMQNTNPESPD-- 173  
Qy 254 HDAVRLNRYKIILLHGGADM-----TKNLAGKPTDVLVOLWQADT 296  
Db 174 -----VTIHAATPQFIIGPGVVNLDTETGVSNAVQFGNSST 209

RESULT 15  
ID GABP HUMAN STANDARD; PRT; 383 AA.  
AC Q06547;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE GA BINDING PROTEIN BETA-1 CHAIN (GABP-BETA-1 SUBUNIT) (TRANSCRIPTION  
DE FACTOR E4TF1-53) (GABPB1) (NUCLEAR RESPIRATORY FACTOR-2 SUBUNIT BETA  
DE 2)  
GN GABPB1 OR GABPB OR E4TF1B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93180783; PubMed=8441384;  
RA Watanabe H., Sawada J.-I., Yano K.-I., Yamaguchi K., Goto M.,  
RA Handa H.;  
RT "cDNA cloning of transcription factor E4TF1 subunits with Ets and  
RT notch motifs";  
RL Mol. Cell. Biol. 13:1385-1391(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95097980; PubMed=7799916;  
RA Gugreja S., Virbasius J.V., Scarpulla R.C.;  
RT "Four structurally distinct, non-DNA-binding subunits of human  
RT nuclear respiratory factor 2 share a conserved transcriptional  
RT activation domain";  
RL Mol. Cell. Biol. 15:102-111(1995).  
CC -1- FUNCTION: TRANSCRIPTION FACTOR CAPABLE OF INTERACTING WITH PURINE  
CC RICH REPEATS (GA REPEATS). NECESSARY FOR THE EXPRESSION OF THE  
CC ADENOVIRUS E4 GENE.  
CC -1- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.

CC -1- ALTERNATIVE PRODUCTS: THE BETA-1 AND BETA-2 SUBUNITS ARE  
CC PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE AND ONLY  
CC DIFFER IN THEIR C-TERMINAL EXTREMITY.  
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; D13317; BAA02574.1; -  
DR EMBL; U13046; AAA65708.1; -  
DR HSP; O00420; IAWC.  
DR TRANSFAC; T01391; -  
DR MIM; 600610; -  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF00023; ank; 3.  
DR SMART; SM00248; ANK; 3.  
DR PROSITE; PS50088; ANK\_REPEAT; 3.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
KW Transcription regulation; Nuclear protein; Alternative splicing;  
KW ANK repeat; Repeat. 34 ANK 1.  
FT REPEAT 37 66 ANK 2.  
FT REPEAT 70 99 ANK 3.  
FT REPEAT 103 132 ANK 4.  
FT REPEAT 136 166 ANK 5.  
SQ SEQUENCE 383 AA; 41321 MW; 07E7081A60016288 CRC64;

Query Match 12.3%; Score 207.5; DB 1; Length 383;  
Best Local Similarity 27.8%; Pred. No. 1.9e-07;  
Matches 64; Conservative 34; Mismatches 101; Indels 31; Gaps 5;

Qy 75 IIDVGGIONLIELKRRKQKRDALAAASHEPPEPEETGPVDETFLLKAAVEGKMKVIE 134  
Db 3 LVDLG--KKLLEAARAGQDDEVIRILMANGAPFTTDWLGTSPH-----LAAQYGHYSTTE 55  
Qy 135 KFLADGGSADTCQDPRRTALHRASLEGHMEILEKLLDNGATVDFODRLDCTAMHWACRG 194  
Db 56 VLLRAGVSRDARTKVDRTPLHMAASEGHASIVEVLKKGADVNKADMLKMTALHWATEHN 115  
Qy 195 HLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIYEHF-LSGLGLEINARDREGDTAL 253  
Db 116 HOEVVELLIKYGADVHTQSKFCKTAFDISIDNGNEDLAEILQIAMQNTNPESPD-- 173  
Qy 254 HDAVRLNRYKIILLHGGADM-----TKNLAGKPTDVLVOLWQADT 296  
Db 174 -----VTIHAATPQFIIGPGVVNLDTETGVSNAVQFGNSST 209

Search completed: April 17, 2002, 08:13:23  
Job time: 181 sec

THIS PAGE BLANK (USPTO)

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 17, 2002, 08:09:42 ; Search time 12.56 Seconds  
(without alignments)  
589.457 Million cell updates/sec

Title: US-09-758-593A-1

Perfect score: 1693

Sequence: 1 MEDSAVORATALIEQRLAQ.....NGLEGPNDSGRETPQVPQAQ 329

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	249.5	14.7	949	4	US-09-196-387-10
2	249.5	14.7	1327	4	US-09-196-387-2
3	234.5	13.9	1839	2	US-09-172-977-4
4	231.5	13.7	843	2	US-09-172-977-3
5	228	13.5	1088	4	US-09-082-059-2
6	224	13.2	679	2	US-08-462-481-4
7	224	13.2	679	2	US-08-436-771-6
8	224	13.2	679	2	US-08-434-998-6
9	224	13.2	679	2	US-08-487-797-6
10	224	13.2	679	2	US-08-701-005A-4
11	224	13.2	679	2	US-08-479-895-4
12	224	13.2	679	5	PCT-US95-02058-6
13	223	13.2	422	3	US-08-484-575A-6
14	223	13.2	422	3	US-08-477-459-6
15	223	13.2	422	3	US-08-479-869-6
16	223	13.2	422	4	US-08-486-414-6
17	223	13.2	422	5	PCT-US94-01826A-6
18	223	13.2	422	5	PCT-US94-02252A-6
19	219	12.9	679	3	US-08-943-956A-4
20	219	12.9	835	4	US-09-291-839-2
21	218.5	12.9	673	4	US-09-196-387-8
22	218.5	12.9	741	3	US-08-943-956A-2
23	217	12.8	741	2	US-08-462-481-2
24	217	12.8	741	2	US-08-436-771-2
25	217	12.8	741	2	US-08-436-771-4
26	217	12.8	741	2	US-08-434-998-2
27	217	12.8	741	2	US-08-434-998-4

28	217	12.8	741	2	US-08-487-797-2	Sequence 2, Appli
29	217	12.8	741	2	US-08-487-797-4	Sequence 4, Appli
30	217	12.8	741	2	US-08-701-005A-2	Sequence 2, Appli
31	217	12.8	741	2	US-08-479-895-2	Sequence 2, Appli
32	217	12.8	741	5	PCT-US95-02058-2	Sequence 2, Appli
33	217	12.8	741	5	PCT-US95-02058-4	Sequence 4, Appli
34	205.5	12.1	1745	2	US-09-031-485-33	Sequence 33, Appli
35	205.5	12.1	1745	2	US-08-847-429A-33	Sequence 33, Appli
36	205.5	12.1	1745	3	US-09-065-474-33	Sequence 33, Appli
37	203.5	12.0	1423	4	US-08-810-712-10	Sequence 10, Appli
38	202.5	12.0	314	2	US-08-989-478-3	Sequence 3, Appli
39	202.5	12.0	314	3	US-08-986-685-3	Sequence 3, Appli
40	201	11.9	352	3	US-09-065-474-139	Sequence 139, App
41	198	11.7	191	2	US-09-031-485-20	Sequence 20, Appli
42	198	11.7	191	2	US-08-847-429A-20	Sequence 20, Appli
43	198	11.7	191	3	US-09-065-474-20	Sequence 20, Appli
44	198	11.7	314	2	US-08-989-478-4	Sequence 4, Appli
45	198	11.7	314	3	US-08-996-685-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-196-387-10  
; Sequence 10, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; City: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196.387  
; FILING DATE: June 10, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION NUMBER: 09/095,225  
; FILING DATE: June 10, 1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 949 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-196-387-10

Query Match 14.7%; Score 249.5; DB 4; Length 949;  
Best Local Similarity 26.2%; Pred. No. 3.4e-16;  
Matches 91; Conservative 33; Mismatches 110; Indels 113; Gaps 9;

QY 40 LVLEDEKH---HGAQSAALQKVGOERVKTSLDLRRRIIDVGGION-----LIELR 88  
 Db 484 LTYEFKSHLLQAAREADLAKV-----KT---LALRIINFKQPOSHETALHCAVASLH 534  
 QY 89 KKRKO-----KKRDALAAASHPEPPPEETITGPVDEETFLKAAVEGKMKVIE 134  
 Db 535 PKRKQVTELLLRKGANVNEKNKDFMTPLH-----VAAERAHNDVME 575  
 QY 135 KFLADGGSADTCDFRRRTALHRASLEGHMEILEKLLDNG-----173  
 Db 576 VLHKHGAKMNAIDTIGQTAALHRAALAGHLQTCRLLLSYGSDPSIISLOGFTAAQMGNEAV 635  
 QY 174 -----ATVDFQ-----DRLDC-----TAMHWACRG 194  
 Db 636 QQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYN 695  
 QY 195 HLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIYEHFSLGLEINARDREGDTALH 254  
 Db 696 RVSVEYLLHHGADVHAKDKGLVPLHNACSYGHVEVAELLVRHGASVNVADLWKFTPLH 755  
 QY 255 DAVRLNRYKIILKLLHGDMMTKNLAGTPTDVLQWQADTRHALE 301  
 Db 756 EAAAGKYEICKLLKHGADPTKKNRDNTPDLVKEGDTDIQDLK 802

RESULT 2

US-09-196-387-2  
 ; Sequence 2, Application US/09196387  
 ; Patent No. 6277613  
 ; GENERAL INFORMATION:  
 ; APPLICANT: de Lange, Titia  
 ; APPLICANT: Smith, Susan  
 ; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
 ; TITLE OF INVENTION: OF USE THEREOF  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Klauber & Jackson  
 ; STREET: 411 Hackensack Avenue, 4th Floor  
 ; CITY: Hackensack  
 ; STATE: New Jersey  
 ; COUNTRY: USA  
 ; ZIP: 07601  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/196,387  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/095,225  
 ; FILING DATE: June 10, 1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Jackson Esq., David A.  
 ; REGISTRATION NUMBER: 26,742  
 ; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 201-487-5800  
 ; TELEFAX: 201-343-1684  
 ; TELEX: 133521  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1327 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; HYPOTHETICAL: NO  
 ; US-09-196-387-2

Query Match 14.7%; Score 249.5; DB 4; Length 1327;  
 Best Local Similarity 26.2%; Pred. No. 5.6e-16;  
 Matches 91; Conservative 33; Mismatches 110; Indels 113; Gaps 9;  
 QY 40 LVLEDEKH---HGAQSAALQKVGOERVKTSLDLRRRIIDVGGION-----LIELR 88  
 Db 484 LTYEFKSHLLQAAREADLAKV-----KT---LALRIINFKQPOSHETALHCAVASLH 534  
 QY 89 KKRKO-----KKRDALAAASHPEPPPEETITGPVDEETFLKAAVEGKMKVIE 134  
 Db 535 PKRKQVTELLLRKGANVNEKNKDFMTPLH-----VAAERAHNDVME 575  
 QY 135 KFLADGGSADTCDFRRRTALHRASLEGHMEILEKLLDNG-----173  
 Db 576 VLHKHGAKMNAIDTIGQTAALHRAALAGHLQTCRLLLSYGSDPSIISLOGFTAAQMGNEAV 635  
 QY 174 -----ATVDFQ-----DRLDC-----TAMHWACRG 194  
 Db 636 QQILSESTPIRTSDVDYRLLEASKAGDLETVKQLCSSQNVNCRDLEGRHSTPLHFAAGYN 695  
 QY 195 HLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIYEHFSLGLEINARDREGDTALH 254  
 Db 696 RVSVEYLLHHGADVHAKDKGLVPLHNACSYGHVEVAELLVRHGASVNVADLWKFTPLH 755  
 QY 255 DAVRLNRYKIILKLLHGDMMTKNLAGTPTDVLQWQADTRHALE 301  
 Db 756 EAAAGKYEICKLLKHGADPTKKNRDNTPDLVKEGDTDIQDLK 802

RESULT 3  
 US-09-172-977-4  
 ; Sequence 4, Application US/09172977  
 ; Patent No. 5989863  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Corley, Neil C.  
 ; APPLICANT: Yue, Henry  
 ; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
 ; FILE REFERENCE: PF-0615 US  
 ; CURRENT APPLICATION NUMBER: US/09/172,977  
 ; CURRENT FILING DATE: 1998-10-14  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 4  
 ; LENGTH: 1839  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE: -  
 ; OTHER INFORMATION: g29491  
 ; US-09-172-977-4

Query Match 13.9%; Score 234.5; DB 2; Length 1839;  
 Best Local Similarity 29.5%; Pred. No. 2.7e-14;  
 Matches 65; Conservative 32; Mismatches 100; Indels 23; Gaps 3;  
 QY 124 AAVEGKMKVIEKFLADGGSADTCDFRRRTALHRASLEGHMEILEKLLDNGATVDFQDRLD 183  
 Db 636 AAKNQMQIASTLLNNGAETNIVTKQVTPHLASQEGHTDMVTLDDKGANIHMSTKSG 695  
 QY 184 CTAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGQVEIYEHFSLGLEIN 243  
 Db 696 LLSLAAQEDKVNADILTHGADQDAHTKLGTYPLIVACHYGNVKNVNFLLKQGANVN 755  
 QY 244 ARDREGDTALHDVRLNRYKIILKLLHGDMMTKNLAGTPTDVLQWQADTRHALE 296  
 Db 756 AKTKNGYPLHOAAQOGHTHIINVLQHGAKPNATANGNTALAKRIGYISVVDTLKV 815  
 QY 297 -----RHALEHPEPGAENHLEGPNDSGRET 322  
 Db 816 VTEEVTTTTTITERKHLNVPMTE--VLDVSDDEGDDT 853

RESULT 4  
US-09-172-977-3  
; Sequence 3, Application US/09172977  
; Patent No. 5989863  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
; FILE REFERENCE: PF-0615 US  
; CURRENT APPLICATION NUMBER: US/09/172.977  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 3  
; LENGTH: 843  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE: -  
; OTHER INFORMATION: g1841966  
US-09-172-977-3

Query Match 13.7%; Score 231.5; DB 2; Length 843;  
Best Local Similarity 29.1%; Pred. No. 1.8e-14;  
Matches 64; Conservative 33; Mismatches 100; Indels 23; Gaps 3;  
QY 124 AAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQDRLD 183  
DB 601 AAKNMQIASTPLNYGAETNTVTKOGYTPPLHLASQEGHDMVTLVLEKGANIHMSTKSG 660  
QY 184 CTAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGOVEIVEHFLSLGLEIN 243  
DB 661 LTLHLAAEEDKVNADILTRGADQDAYTKLGYTPLIVACHYGNVKNVFLKOGANVN 720  
QY 244 ARDREGDTALHDVARNRYKIIKLLHGGADMMTKNLAKGTPDLVQ----LWQADT--- 296  
DB 721 AKTKNGYTPLHQAQGGHTIIINVLQHGAKPNATTANGNTALATKRLGIVSVVDTLKV 780  
QY 297 -----RHLEHPEPGEAHNGLEGNDSGRET 322  
DB 781 VTEVTTTTTITEKHKLNAPEMTTE--VLDVSDEEGDDT 818

RESULT 5  
US-09-082-059-2  
; Sequence 2, Application US/09082059A  
; Patent No. 6225086  
; GENERAL INFORMATION:  
; APPLICANT: Moritow, Jon S.  
; APPLICANT: Devarajan, Prasad  
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification  
; FILE REFERENCE: 44574-5002-US  
; CURRENT APPLICATION NUMBER: US/09/082.059A  
; CURRENT FILING DATE: 1998-05-21  
; EARLIER APPLICATION NUMBER: 60/047356  
; EARLIER FILING DATE: 1997-05-21  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1088  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-082-059-2

Query Match 13.5%; Score 228; DB 4; Length 1088;  
Best Local Similarity 33.3%; Pred. No. 5.7e-14;  
Matches 54; Conservative 26; Mismatches 82; Indels 0; Gaps 0;  
QY 124 AAVEGKMKVIEKFLADGGSADTCDFRRTALHRASLEGHMEILEKLLDNGATVDFQDRLD 183

DB 28 ACKNRIKVMELLKKGASIQAVTESGLTPIHVAAFMGHVNIVSQLMHHGASPNNTTVRG 87  
QY 184 CTAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLHVAVRTGOVEIVEHFLSLGLEIN 243  
DB 88 ETALHMAARSGQAEVRYLVLDQGAQVEAKAKDDQPLHLISARLGRADIVQQLQOGASPN 147  
QY 244 ARDREGDTALHDVARNRYKIIKLLHGGADMMTKNLAKGTP 285  
DB 148 AATTSYTPLHLSAREGHEDVAAFLLDHGASLSITTKGFTP 189

RESULT 6  
US-08-462-481-4  
; Sequence 4, Application US/08462481  
; Patent No. 5840577  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Hassel, Bret A.  
; APPLICANT: Zhou, Almin  
; TITLE OF INVENTION: Animal 2-5A- Dependent RNases and  
; TITLE OF INVENTION: Encoding Sequence therefor  
; Patent No. 5840577  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
; ADDRESSEE: Russell  
; STREET: 200 East Broward Boulevard  
; CITY: Fort Lauderdale  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 33301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462.481  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/028,086  
; FILING DATE: 08-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Manso, Peter J  
; REGISTRATION NUMBER: 32,264  
; REFERENCE/DOCKET NUMBER: CL11363-13  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 305/527/2498  
; TELEFAX: 305/764/4996  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 679 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-462-481-4

Query Match 13.2%; Score 224; DB 2; Length 679;  
Best Local Similarity 28.1%; Pred. No. 7.3e-14;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;  
QY 41 VLEDEKHGHAQSAALQKVKQGERVKTSLDLRREIDVCGIONLIE-----LRKKRQ 93  
DB 22 VVEDD-----SLIKAVQKGDVVRV-----QQLLEKADANACEDTGWTPPLHNAVQA 69  
QY 94 KKRDA--LAASHEPPEPEITGPVDEETFLKAAVEGKMKVIEKFLADGGSADTCDO--- 148  
DB 70 GRVDIVNLLSHGADPHRRKKNGAT---PFIAGIQGDVKLLLEILSCGADVNECDENGF 126  
QY 149 -----FRR-----TALHRASLEGHMEILEK 168

Db 127 TAFMEAAERGAELRFLFAKGANVNLRRQTTKDRRLKQGGATALMSAAERGHLEVLRI 186  
QY 169 LL-DNGATVDFODR-----LDCATAMHWACRGHLEVVKLLQSHGADTVNRDKLLSTPLHV 222  
Db 187 LNDMKAEDVARDNMGRLNIRTLNWDCEVNE-EITSILIOHGADVNRGERTPLIA 245  
QY 223 AVRTGQVEIVEHFLSL-GLEINARDREGDTALHDVRLNRYKIIKLLHGHGADMMTKNLA 281  
Db 246 AVERKHTGLVQMLLSREGINDARDNEGTALLIAVDKOLKEIVQLLEKGDAC----- 300  
QY 282 GKTPTDLVOLMOADTRH 298  
Db 301 ----DDLVL-WIARRNH 311

## RESULT 7

US-08-436-771-6  
; Sequence 6, Application US/08436771  
; Patent No. 5861300  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Sengupta, Dibyendu N.  
; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
; ADDRESSEE: Russell  
; STREET: 200 E. Broward Boulevard  
; CITY: Fort Lauderdale  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 33301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/436,771  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/198,973  
; FILING DATE: 18-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Manso, Peter J.  
; REGISTRATION NUMBER: 32,264  
; REFERENCE/DOCKET NUMBER: CL11363-16  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 305/527/2498  
; TELEFAX: 305/764/4996  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 679 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-436-771-6

Query Match 13.2%; Score 224; DB 2; Length 679;  
Best Local Similarity 28.1%; Pred. No. 7.3e-14;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;  
QY 41 VLEDEKHGAQSAALQKVGQERVRTSLDLREIIDVGGIQLNIE-----LRKKRQ 93  
Db 22 VVEDD-----SSLKAVQKGVVRY-----QOLLEKADANACEDTWGTPPLHNAVOA 69  
QY 94 KKRD--LAASHEPPPEPEITGPVDEETFLKAAVGGKMKVIEKFLADGSGADTCDO--- 148  
Db 70 GRVDIYNLLSHGADPHRRKNGAT---PFITAGIOGVKLLLEILLSCGADVNECDENG 126

QY 149 -----FRR-----TALHRASLEIGHMEILEK 168  
Db 127 TAFMEAAERGAELRFLFAKGANVNLRRQTTKDRRLKQGGATALMSAAERGHLEVLRI 186  
QY 169 LL-DNGATVDFODR-----LDCATAMHWACRGHLEVVKLLQSHGADTVNRDKLLSTPLHV 222  
Db 187 LNDMKAEDVARDNMGRLNIRTLNWDCEVNE-EITSILIOHGADVNRGERTPLIA 245  
QY 223 AVRTGQVEIVEHFLSL-GLEINARDREGDTALHDVRLNRYKIIKLLHGHGADMMTKNLA 281  
Db 246 AVERKHTGLVQMLLSREGINDARDNEGTALLIAVDKOLKEIVQLLEKGDAC----- 300  
QY 282 GKTPTDLVOLMOADTRH 298  
Db 301 ----DDLVL-WIARRNH 311

## RESULT 8

US-08-434-998-6  
; Sequence 6, Application US/08434998  
; Patent No. 5866781  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Sengupta, Dibyendu N.  
; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
; TITLE OF INVENTION: Cells and Methods  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
; ADDRESSEE: Russell  
; STREET: 200 E. Broward Boulevard  
; CITY: Fort Lauderdale  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 33301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/434,998  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/198,973  
; FILING DATE: 18-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Manso, Peter J.  
; REGISTRATION NUMBER: 32,264  
; REFERENCE/DOCKET NUMBER: CL11363-16  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 305/527/2498  
; TELEFAX: 305/764/4996  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 679 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-434-998-6

Query Match 13.2%; Score 224; DB 2; Length 679;  
Best Local Similarity 28.1%; Pred. No. 7.3e-14;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;  
QY 41 VLEDEKHGAQSAALQKVGQERVRTSLDLREIIDVGGIQLNIE-----LRKKRQ 93  
Db 22 VVEDD-----SSLKAVQKGVVRY-----QOLLEKADANACEDTWGTPPLHNAVOA 69  
QY 94 KKRD--LAASHEPPPEPEITGPVDEETFLKAAVGGKMKVIEKFLADGSGADTCDO--- 148

DB 70 GRVDIVNLLSHGADPHRRKKGAT---PFIAGIQGDVKLEILLSCGADVNECDNGF 126  
QY 149 -----FRR-----TALHRASLEGHMEILEK 168  
DB 127 TAFMEAERGNAELRFLFAKGANVNLRRQTTKDKRRLKQGGATALMSAAEKGHLEVLRI 186  
QY 169 LL-DNGATVDFQDR-----LDCATAMHWACRGHLEVVKLLQSHGADTVNRDKLLSTPLHV 222  
DB 187 LNDMKAEDVARDNMGRLNLTLLNWDCEVNE-BITSILIOHGADVNVVRGKGTPLIA 245  
QY 223 AVRTQGVETVEHFLSL-GLEINARDREGDTALHDVRLNRKYIKILLHGHGADMTKNLA 281  
DB 246 AVERKHTGLVQMLLSREGINIDARDNEGTALLIAVDKOLKEIVQLLLEKGDAC----- 300  
QY 282 GKTPDVLVOLWQADTRH 298  
DB 301 ----DDL-V-WIARRNH 311

## RESULT 9

US-08-487-797-6  
; Sequence 6, Application US/08487797  
; Patent No. 5866787  
; GENERAL INFORMATION:

APPLICANT: Silverman, Robert H.  
APPLICANT: Sengupta, Dibyendu N.  
TITLE OF INVENTION: Transgenic Plants Co-Expressing A  
TITLE OF INVENTION: Functional Human 2-5A System  
NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Holland & Knight  
STREET: One E. Broward Boulevard, #1300  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,797  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL1363-16(C)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/468-7811  
TELEFAX: 305/463-2030

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 679 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-487-797-6

Query Match 13.2%; Score 224; DB 2; Length 679;  
Best Local Similarity 28.1%; Pred. No. 7.3e-14;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;

QY 41 VLEDEKHHGAQAALQKVKQGERVKTSLDLRREIIDVGGIONLIE-----LRKKRKQ 93  
DB 22 VVEDD-----SSLKAVQKGVVRY-----QQLLEKADANACEDTGWTPPLHNAVOA 69  
QY 94 KKRA-D-LAASHEPPEPEITGPVDEETFLKAAVEGKMKVTEKFLADGSGADTCDQ--- 148  
DB 70 GRVDIVNLLSHGADPHRRKKGAT---PFIAGIQGDVKLEILLSCGADVNECDNGF 126

QY 149 -----FRR-----TALHRASLEGHMEILEK 168  
DB 127 TAFMEAERGNAELRFLFAKGANVNLRRQTTKDKRRLKQGGATALMSAAEKGHLEVLRI 186  
QY 169 LL-DNGATVDFQDR-----LDCATAMHWACRGHLEVVKLLQSHGADTVNRDKLLSTPLHV 222  
DB 187 LNDMKAEDVARDNMGRLNLTLLNWDCEVNE-BITSILIOHGADVNVVRGKGTPLIA 245  
QY 223 AVRTQGVETVEHFLSL-GLEINARDREGDTALHDVRLNRKYIKILLHGHGADMTKNLA 281  
DB 246 AVERKHTGLVQMLLSREGINIDARDNEGTALLIAVDKOLKEIVQLLLEKGDAC----- 300  
QY 282 GKTPDVLVOLWQADTRH 298  
DB 301 ----DDL-V-WIARRNH 311

## RESULT 10

US-08-701-005A-4  
; Sequence 4, Application US/08701005A  
; Patent No. 5877019  
; GENERAL INFORMATION:

APPLICANT: Silverman, Robert H.  
APPLICANT: Hassel, Bret A.

APPLICANT: Zhou, Almin  
TITLE OF INVENTION: Animal 2-5A-Dependent RNases and  
TITLE OF INVENTION: Encoding Sequence Therefor  
Patent No. 5877019

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rothwell, Figg, Ernst & Kurz  
STREET: 555 - 13th Street, N.W., Suite 701 E  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/701,005A  
FILING DATE: 21-AUG-1996  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/141,304  
FILING DATE: 22-OCT-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/028,086  
FILING DATE: 08-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Jondle, Robert J.

REGISTRATION NUMBER: 33,915

REFERENCE/DOCKET NUMBER: N1255-111

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-783-6040

TELEFAX: 202-783-6031

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 679 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-701-005A-4

Query Match 13.2%; Score 224; DB 2; Length 679;  
Best Local Similarity 28.1%; Pred. No. 7.3e-14;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;  
QY 41 VLEDEKHHGAQAALQKVKQGERVKTSLDLRREIIDVGGIONLIE-----LRKKRKQ 93

Db 22 VVEDD-----SSLIKAVQGDVVRV-----QOLLEKAGADANACEDTWGTPPLHNAVQA 69  
QY 94 KKRDAA--LAASHEPPPEEITGPDVEETFLKAAVEGKMKVIEKFLADGGSADTCDO--- 148  
Db 70 GRDIVNLLSHGADPHRRKNGAT---PFIAGIQGDVKLLLELLSCGADVNECDNGF 126  
QY 149 -----FRR-----TALHRASLEGHMEILEK 168  
Db 127 TAFMEAAERGAELRFLFAGKANVNLRRQTTKDKRRLKOGGATALMSAAEKHLEVLRI 186  
QY 169 LL-DNGATVDFODR-----LDCTAMHWACRGHLEVVVKKLQSHGADTVNRDKLLSTPLHV 222  
Db 187 LNDKMAEVDARDMGNRNALIRTLNWDCEVE-EITSILIQHGADVNVVRGERGKTPLIA 245  
QY 223 AVRTGOVEIVEHFLSL-GLEINARDREGDTALHDVRLNRYKIIKILLHGGADMTKNLA 281  
Db 246 AVERKHTGLVQMLLSREGINARDNEGKTALLIADVQKQLEIVQLLEKGAADK----- 300  
QY 282 GKPTDVLVQWQADTRH 298  
Db 301 ----DDLW--WIARRNH 311

RESULT 11  
US-08-479-895-4  
; Sequence 4, Application US/08479895  
; Patent No. 5972678  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Hassel, Bret A.  
; APPLICANT: Zhou, Aimin  
; TITLE OF INVENTION: Animal 2-5A- Dependent RNases and  
; TITLE OF INVENTION: Encoding Sequence Therefor  
; Patent No. 5972678  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
; ADDRESSEE: Russell  
; STREET: 200 East Broward Boulevard  
; CITY: Fort Lauderdale  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 33301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/479,895  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/028,086  
; FILING DATE: 08-MAR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Manso, Peter J.  
; REGISTRATION NUMBER: 32,264  
; REFERENCE/DOCKET NUMBER: CL11363-13  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 305/527/2498  
; TELEFAX: 305/764/4996  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 679 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-479-895-4

Query Match 13.2% Score 224; DB 2; Length 679;

Best Local Similarity 28.1%; Pred. No. 7 3e-14;  
Matches 89; Conservative 41; Mismatches 101; Indels 86; Gaps 13;  
QY 41 VLEDEKHGAQAALQKVKQERKRTSLDLREIIVDVGIIQNLIIE-----LRKKRQ 93  
Db 22 VVEDD-----SSLIKAVQGDVVRV-----QOLLEKAGADANACEDTWGTPPLHNAVQA 69  
QY 94 KKRDAA--LAASHEPPPEEITGPDVEETFLKAAVEGKMKVIEKFLADGGSADTCDO--- 148  
Db 70 GRDIVNLLSHGADPHRRKNGAT---PFIAGIQGDVKLLLELLSCGADVNECDNGF 126  
QY 149 -----FRR-----TALHRASLEGHMEILEK 168  
Db 127 TAFMEAAERGAELRFLFAGKANVNLRRQTTKDKRRLKOGGATALMSAAEKHLEVLRI 186  
QY 169 LL-DNGATVDFODR-----LDCTAMHWACRGHLEVVVKKLQSHGADTVNRDKLLSTPLHV 222  
Db 187 LNDKMAEVDARDMGNRNALIRTLNWDCEVE-EITSILIQHGADVNVVRGERGKTPLIA 245  
QY 223 AVRTGOVEIVEHFLSL-GLEINARDREGDTALHDVRLNRYKIIKILLHGGADMTKNLA 281  
Db 246 AVERKHTGLVQMLLSREGINARDNEGKTALLIADVQKQLEIVQLLEKGAADK----- 300  
QY 282 GKPTDVLVQWQADTRH 298  
Db 301 ----DDLW--WIARRNH 311

RESULT 12  
PCT-US95-02058-6  
; Sequence 6, Application PC/TUS9502058  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: SenGupta, Dibyendu N.  
; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
; TITLE OF INVENTION: Cells and Methods  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
; ADDRESSEE: Russell  
; STREET: 200 E. Broward Boulevard  
; CITY: Fort Lauderdale  
; STATE: Florida  
; COUNTRY: USA  
; ZIP: 33301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02058  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/198,973  
; FILING DATE: 18-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Manso, Peter J.  
; REGISTRATION NUMBER: 32,264  
; REFERENCE/DOCKET NUMBER: CL11363-16  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 305/527/2498  
; TELEFAX: 305/764/4996  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 679 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-02058-6



Query Match	13.28;	Score	224;	DB	5;	Length	679;
Best Local Similarity	28.18;	Pred. No.	7.3e-14;				
Matches	89;	Conservative	41;	Mismatches	101;	Indels	86;
						Gaps	13;
QY	41	VLEDEKHGGAQSAALQKVQGERVKYTSLDLRREIDVGGIQNLIE-----LRKKRKQ	93				
Db	22	VVEDD-----SSLIKAVQGVVRV-----QOLLEKGADANACEDTGWGTPPLHNAVQA	69				
QY	94	KRDA--LAASHEPPPEETGPVDEETFLLKAAVEGKMVKVIERFLADGGSADTCQ---	148				
Db	70	GRVDIVNLLSHGADPHRRKKNGAT--PFIAGIQGDVKLEILLSCGADVNECEDNGF	126				
QY	149	-----PFR-----TALHRASLEGHMEYLEK	168				
Db	127	TAFMEAAERGNAAELFLFAGKANVNRQTTDKRRLLKOGGATALMSAAEKHGLEVLR	186				
QY	169	LL-DNGATVDFODR-----LDCTAMHWACRGHLEVVKLLQSHGADTVNRDKLLSTPLHV	222				
Db	187	LLNDMKAEVDARDNMGRRNALIRTLNDWCENVE-EITSILIQHGADVNRGERGKTPLIA	245				
QY	223	AVRTGQVEYHEHFLSL-GLEINARDREGDTTALHDVRLNRYKIITKLILLGADMMTKNLA	281				
Db	246	AVERKHTGLVQMLLSREGINIDARNEGKTALLTAVDKQLAEIVOLLLEKGADRC----	300				
QY	282	GKPTDVLQVLQWADTRH	298				
Db	301	----DDLV--WIARRNH	311				

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match 13.2%; Score 223; DB 2; Length 422;  
Best Local Similarity 25.9%; Pred. No. 4.6e-14;  
Matches 67; Conservative 48; Mismatches 98; Indels 46; Gaps 7;

QY 68 SLDURREII--DVGGIONLTIELRKKRKQKRDALAAASHEPPPEEITGPDVEETFLKAA 125  
| : | | | | : : : | | : | : | : | : | : | : |

QY 223 AVRTGQVEIVHFSLGLEINARDREGDTALHDVRLNRYKIIKLLLLHGADMMTKNLG 282  
 Db 164 ATRNSYEIIKLLLEKAYANVKNYGNPSPLHNAKYGDYACIKLVLDHTNINISKNCNG 223  
 QY 283 KPTPTDLVQLWQADTRHALE 301  
 Db 224 VTPLHNALLY---NRSAGE 239

QY 283 KPTPTDLVQLWQADTRHALE 301  
 Db 224 VTPLHNALLY---NRSAGE 239

Search completed: April 17, 2002, 08:10:35  
 Job time: 53 sec

RESULT 15  
 US-08-479-869-6  
 ; Sequence 6, Application US/08479869  
 ; Patent No. 6123949  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cochran Ph.D, Mark D  
 ; TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043 and  
 ; TITLE OF INVENTION: Uses Thereof  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: John P. White  
 ; STREET: 30 Rockefeller Plaza  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10112  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/479,869  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/024,156  
 ; FILING DATE: 26-FEB-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White Esq, John P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)977-9550  
 ; TELEFAX: (212)664-0525  
 ; TELEX: 422523  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 422 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; -08-479-869-6

Query Match 13.2%; Score 223; DB 3; Length 422;  
 Best Local Similarity 25.9%; Pred. No. 4.6e-14;  
 Matches 67; Conservative 48; Mismatches 98; Indels 46; Gaps 7;  
 QY 68 SLDLRREII--DVGGTONLIELKRRKQKRDALAAASHPEPEPEEITGPVDEETFLKAA 125  
 Db 2 SVDWRTIYSGDISLVEKLI-----KNGNCINIS-----VEETTTPL-----IDAI 43  
 QY 126 VEGKMKVIEKFLADGGSADTCD-QFRRTALHRASLEGH----- 162  
 Db 44 RTGNKIVELFKHGAQVNHVTKIPNPLLTAKIGSHDIVKLLLLINGVDTSILPVCIN 103  
 QY 163 MEILEKLLDNGATVDFQDELCDTAMHWACRGHLEVVKLLQSHGADTNVRDKLLSTPLHV 222  
 Db 104 KEMIKTILDSGVKVKNTAKSKTFLHYAKNNKNDLEVIKMLFEYGADVNIKDDNICYSIHI 163  
 QY 223 AVRTGQVEIVHFSLGLEINARDREGDTALHDVRLNRYKIIKLLLLHGADMMTKNLG 282  
 Db 164 ATRNSYEIIKLLLEKAYANVKNYGNPSPLHNAKYGDYACIKLVLDHTNINISKNCNG 223

---

THIS PAGE BLANK (USPTO)

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 08:09:16 ; Search time 88.05 Seconds  
(without alignments)  
2978.552 Million cell updates/sec

Title: US-09-758-593A-2  
Perfect score: 1158  
Sequence: 1 cagctcgaggagcgaccac.....ctaccacaataaaaagctg 1158

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.2	6.7	1288	US-09-172-977-2	Sequence 2, Appli
2	75.6	6.5	3454	US-09-082-059-1	Sequence 1, Appli
3	64.2	5.5	1789	US-09-035-706-1	Sequence 1, Appli
4	64.2	5.5	1789	US-08-955-841-1	Sequence 3, Appli
5	64.2	5.5	1789	US-09-428-219-3	Sequence 3, Appli
6	62.4	5.4	16442	US-08-781-891-208	Sequence 208, App
7	62.4	5.4	4134	US-09-196-387-1	Sequence 1, Appli
8	62.4	5.4	4491	US-09-196-387-7	Sequence 7, Appli
9	61.8	5.3	2928	US-08-462-481-1	Sequence 1, Appli
10	61.8	5.3	2928	US-08-436-771-1	Sequence 1, Appli
11	61.8	5.3	2928	US-08-436-771-3	Sequence 3, Appli
12	61.8	5.3	2928	US-08-434-998-1	Sequence 1, Appli
13	61.8	5.3	2928	US-08-434-998-3	Sequence 3, Appli
14	61.8	5.3	2928	US-08-487-797-1	Sequence 1, Appli
15	61.8	5.3	2928	US-08-487-797-3	Sequence 3, Appli
16	61.8	5.3	2928	US-08-701-005A-1	Sequence 1, Appli
17	61.8	5.3	2928	US-08-479-895-1	Sequence 1, Appli
18	61.8	5.3	2928	PCT-US95-02058-1	Sequence 1, Appli
19	61.8	5.3	2928	PCT-US95-02058-3	Sequence 3, Appli
20	61.8	5.3	2931	US-08-943-956A-1	Sequence 1, Appli
21	61.2	5.3	4935	US-08-631-097-3	Sequence 3, Appli
22	61.2	5.3	5886	US-08-810-712-9	Sequence 9, Appli
23	60.6	5.2	3516	US-09-188-930-257	Sequence 257, App
24	59.4	5.1	4657	US-09-196-387-9	Sequence 9, Appli
25	59.4	5.1	7218	US-08-232-463-14	Sequence 14, Appli
26	58.6	5.1	2580	US-09-050-863-2	Sequence 2, Appli
27	58.6	5.1	5452	US-09-130-114-1	Sequence 1, Appli

RESULT 1  
US-09-172-977-2  
; Sequence 2, Application US/09172977  
; Patent No. 5989863  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Guegler, Karl J.  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Yue, Henry  
; TITLE OF INVENTION: HUMAN ANKYRIN FAMILY PROTEIN  
; FILE REFERENCE: PF-0615 US  
; CURRENT APPLICATION NUMBER: US/09/172,977  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PERL Program  
; SEQ ID NO 2  
; LENGTH: 1288  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE: -  
; OTHER INFORMATION: 1808075  
US-09-172-977-2

ALIGNMENTS

Query Match	Best Local Similarity	Score	DB 2;	Length	Indels	Gaps
Matches 217;	Conservative	0;	Mismatches 233;	0;	0;	0;
QY 435	acgggggggtcagccgacacgtgcgaccagtccgtcgacagcactgcaccgagcttccc	494				
DB 589	acgaggtgtcagctctgcggccaccctcctgctccatccaccagctcgacgac	648				
QY 495	tggaggccacatgaaatcctggagaagcttctagataatggggccactgtgacattcc	554				
DB 649	agggggagctggaccagctggaaggagcatttcggaaggtgacacccctcgcaacagc	708				
QY 555	aggatcgctgactgcacagcactgattggcctccgcggggccacttagaggtgg	614				
DB 709	cagacagcgcggttcacccccctcattctggcctccgcttggagattgagacgg	768				
QY 615	tgaacctctgcaaacccatggagcagacacccaatgtgaggataagctgtgagacccc	674				
DB 769	ttcgtctctgtggagtggtggtgcgacccccacatcctggcaaaagcgagagagcg	828				
QY 675	cgtgcacgtgagctccggacagggcaggttgagattgtgagacatttctatccctgg	734				
DB 829	cctgtcgctggccagcagcgctcacagacattgtgggtgctgtgtggagcgtg	888				
QY 735	gctggaatcaatgccagagacaggggagatactccctgcatgacgtgtgagcg	794				

Db 889 acgtggacatcaacatctatgattggaatggaggagccactgctgtacgtgtgctg 948  
QY 795 tcaaccgtacaaaatcatcaactgctctcctgcatgggactgacatgatgacaaaga 854  
Db 949 ggaaccacatgaatgctgtggtggtcctgctggtccgagggcgtgacccacacagag 1008  
QY 855 acttggcaggaagaccccgacggactgg 884  
Db 1009 ccgactctgctacaccccgatggaccttg 1038

RESULT 2  
US-09-082-059-1  
; Sequence 1, Application US/09082059A  
; Patent No. 6225086  
; GENERAL INFORMATION:  
; APPLICANT: Morrow, Jon S.  
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification  
; FILE REFERENCE: 44574-5002-US  
; CURRENT APPLICATION NUMBER: US/09/082,059A  
; CURRENT FILING DATE: 1998-05-21  
; EARLIER APPLICATION NUMBER: 60/047356  
; EARLIER FILING DATE: 1997-05-21  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3454  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: (131)..(3394)  
; OTHER INFORMATION: DNA/protein segment- human kidney Ankyrin G119  
US-09-082-059-1

Query Match 6.5%; Score 75.6; DB 4; Length 3454;  
Best Local Similarity 47.5%; Pred. No. 8.2e-10;  
Matches 225; Conservative 0; Mismatches 249; Indels 0; Gaps 0;  
QY 378 ccttctgaagctgcgtgaggagaaatgaagtcattgagaagttcctggctgacg 437  
Db 993 ccgtccatctcagctcaggaagggcagctgacatggtcgtcgtcctcgtgtagaa 1052  
QY 438 ggggttcagccagcagctgcgaccagttcctcgtgacagcactgcacgagcttccctgg 497  
Db 1053 atgcgaatgtaacctgagcaataagagcggcgtgacccactccatttgctgtctcaag 1112  
QY 498 aagggccacatggaatctctggagaagcttctagataatggggccactgtggacttcacgg 557  
Db 1113 aagatcgagtgatgtggcagaagctcgttaaaccaagggtcctatgtggagcccca 1172  
QY 558 atcgtgactgacagcagcattggtggtcctgcgcgggggcaacttagaggtggtga 617  
Db 1173 caaagatgggacacacacacactgcattggtggtgcactatggaataatcaagattgta 1232  
QY 618 aacttctgaaagccatgagcagacacacaaatgtgaggataagctgctgagcaccgcc 677  
Db 1233 atttctgtccagcattctgcaaaagttaatgccaaaagaagtggtatagccat 1292  
QY 678 tgcacgtgagctccggacagcaggtgagattgtgagcactttctatccctgggccc 737  
Db 1293 tacataagcagcagcagcaggggcatacgcataataaatgtcttaactcagaacaag 1352  
QY 738 tggaaatcaatccagagacaggaagggatactgacctgcatgacgtgtgaggtca 797  
Db 1353 cctcccctaaactcaactgtaattgggaatactgcttggcattggcccgcgctg 1412  
QY 798 accgctacaaaatcatcaaacctgctcctcgtcgtgaggtgggctgacatgatgacca 851  
Db 1413 gctacatctcagtagtggaacccttgaaatagtagtaccgaagaaccatgacca 1466

RESULT 3  
US-09-035-706-1  
; Sequence 1, Application US/09035706  
; Patent No. 6001622  
; GENERAL INFORMATION:  
; APPLICANT: Dedhar, Shoukat  
; APPLICANT: Hannigan, Greg  
; TITLE OF INVENTION: Integrin-Linked Kinase and  
; TITLE OF INVENTION: its Uses  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bozicevic & Reed, LLP  
; STREET: 285 Hamilton Avenue, Suite 200  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION NUMBER: US/09/035.706  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sherwood, Pamela J  
; REGISTRATION NUMBER: 36,677  
; REFERENCE/DOCKET NUMBER: KIN-2CIP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-327-3400  
; TELEFAX: 650-327-3231  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1789 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-035-706-1

Query Match 5.5%; Score 64.2; DB 3; Length 1789;  
Best Local Similarity 49.8%; Pred. No. 4.5e-07;  
Matches 162; Conservative 0; Mismatches 163; Indels 0; Gaps 0;  
QY 556 gnatcggtgactgcacagccatcattggcctgcgcggggccacttagagtggt 615  
Db 246 GGACCATATGGCTTCTCCCCCTTGACATGGGCGCTGCCGAGAGGCGCTCTGTGTGT 305  
QY 616 gaaactctgcaaacccatggagcagacaccatgtgagggataagctgctgagcacc 675  
Db 306 TCAGATGTTGATCATGGGGGGGCGACGGATCAATGTAATGAACCGTGGGGATGACACCC 365  
QY 676 gctcagctggcagtcgcggacagggcaggtggagattgtggagcactttctatccctgg 735  
Db 366 CCGTGCATCTGGCAGCAGTCATCGACACCGGTGATATTGTACAGAAAGCTATTGCA 425  
QY 736 cctgaaatcaatgcacagagacaggggaagatactgctcctgcatgacgtgtgaggt 795  
Db 426 GCACACATCAATGAGTGAATGAACACGGGAATGCCCCCTGCACATGATCCCTGTTT 485  
QY 796 caaccgtacaaaatcatcaaacctgctcctgcatgggggctgacatgatgaccaaga 855  
Db 486 GGGCCAAGATCAAGTGGCAGAGGACCTGTGTGCAAAATGGGCCCTTGTGACGATCTG 545  
QY 856 cctggcaggaagaccccgacggac 880







STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196.387  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,225  
FILING DATE: June 10, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4491 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6..207  
US-09-196-387-7

Query Match 5.4%; Score 62; DB 4; Length 4491;  
Best Local Similarity 49.7%; Pred. No. 2.2e-06;  
Matches 158; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 567 actgcagcagcattgagtgccgcccgcggggccacattagagtggtgaaactctgc 626  
DB 2872 ATTCAACCCCTCTGCACCTGGCGAGCTATATAACCTGGAGTAGTGAATATCTC 2931  
QY 627 aaagccatgagcagacacccaatgagggataagctgtgagcaccgcgtgcacgtgg 686  
DB 2932 TAGAGCATGGAGCTGATGTTAATGCCAGCAAGGTGGTTAATTCCTCTTCATATG 2991  
QY 687 cagtcggagcagggcaggtggagattgtgagcactttctatccctggcctggaaatca 746  
DB 2992 CGGCATCTTATGGCATGTTGACATAGCGGCTTTATTGATAAAATACACACGTGTGTAA 3051  
QY 747 atgcagagacgggaaggggatactgcccctgcatgacgtgtgaggtcacaacgcctaca 806  
DB 3052 ATGCAACAGATAAGTGGGCGTTTACTCCCTCCATGAAGCAGCCCAAGAAAGGACGC 3111  
QY 807 aatcatcaaatgctgctccctgagggctgacatgatgacaaagacccggcaggaa 866  
DB 3112 ACTGTGCGCCCTCTCTCTAGCGCATGGTGGACAGCCCAACCATGTAAGAAACCAAGAGGCC 3171  
QY 867 agaccccgagcagcctgg 884  
DB 3172 AGACGCTCTGGATCTGG 3189

RESULT 9  
US-08-462-481-1  
; Sequence 1, Application US/08462481  
; Patent No. 5840577  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Hassel, Bret A.

APPLICANT: Zhou, Almin  
TITLE OF INVENTION: Animal 2-5A- Dependent RNases and  
TITLE OF INVENTION: Encoding Sequence Therefor  
Patent No. 5840577  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462.481  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/028,086  
FILING DATE: 08-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-13  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305/527/2498  
TELEFAX: 305/764/4996  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2928 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 104..2326  
US-08-462-481-1

Query Match 5.3%; Score 61.8; DB 2; Length 2928;  
Best Local Similarity 57.8%; Pred. No. 2.1e-06;  
Matches 130; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

QY 616 gaaacttctgcaaacgcatgagcagacacccaatgtgagggataagctgctgagacccc 675  
DB 769 GCATCTGCTGCTGGACCATGGGCTGATGTCATATGAGGGGAGAAAGAGGAAGACTCC 828  
QY 676 gctgcacgtgagcagtcgagcagggcaggtgagagatgtgagcactttctatccctggg 735  
DB 829 CTTGATCTCTGGCAGTGGAGAAAGCACTTGGGTTGGTGCAGAGGCTTCTGGAGCAAGA 888  
QY 736 cc--tggaaatcaatgccagagacaggaaggggatactgcccctgcatgacgctgtgag 792  
DB 889 GCACATAGAGATTAAATGACACAGACAGTATGCAAAACAGCACTGCTGCTGCTGTGA 948  
QY 793 gctcaaccgctcaaaatcatcaaaactgctgctcctgcatggggc 837  
DB 949 ACTCAAACTGAAGAAATCGCCGAGTTGCTGTGCAAAACCTGGAGC 993

RESULT 10  
US-08-436-771-1  
; Sequence 1, Application US/08436771  
; Patent No. 5861300  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Sengupta, Dibyendu N.

;; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
;; TITLE OF INVENTION: Cells and Methods  
;; NUMBER OF SEQUENCES: 11  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Ruten, Barnett, McClosky, Smith, Schuster &  
;; ADDRESSEE: Russell  
;; STREET: 200 E. Broward Boulevard  
;; CITY: Fort Lauderdale  
;; STATE: Florida  
;; COUNTRY: USA  
;; ZIP: 33301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/436,771  
;; FILING DATE:  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/198,973  
;; FILING DATE: 18-FEB-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Manso, Peter J.  
;; REGISTRATION NUMBER: 32,264  
;; REFERENCE/DOCKET NUMBER: CL11363-16  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 305/527/2498  
;; TELEFAX: 305/764/4996  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2928 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 104...2326  
;; US-08-436-771-1

Query Match 5.3%; Score 61.8; DB 2; Length 2928;  
Best Local Similarity 57.8%; Pred. No. 2.1e-06;  
Matches 130; Conservative 0; Mismatches 92; Indels 3; Gaps 1;  
QY 616 gaaactctgcaagccatggacagacacccaatgtgaggataagctgctgagaccccc 675  
Db 769 GCATCTGCTGCTGACCATGGGCTGATGTCATGAGGGGAGAAAGAGGGAAGACTCC 828  
QY 676 gctgcacgtgagctccgagacagggcaggtggagattgtgagcactttctatccctggg 735  
Db 829 CCTGATCTGCGAGTGGAGAGAGCACTTGGGTTGGTGCAGAGGCTTCTGGAGCAAGA 888  
QY 736 cc---tggaataatccagagacaggggaaggatactgcccctgcatgacgtgtgag 792  
Db 889 GCACATAGAGATTAAATGACACAGACAGTGTGATGCAAAACAGACACTGCTGCTGTGTTGA 948  
QY 793 gctcaacgcgtacaaatcaaatcactgctcctcgtcatggggc 837  
Db 949 ACTCAAACTGAAGAAATGCCGAGTTGCTGTGCAAAACGTGGAGC 993

RESULT 11  
US-08-436-771-3  
; Sequence 3, Application US/08436771  
; Patent No. 5861300  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Sengupta, Dibyendu N.  
; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
; TITLE OF INVENTION: Cells and Methods

;; NUMBER OF SEQUENCES: 11  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Ruten, Barnett, McClosky, Smith, Schuster &  
;; ADDRESSEE: Russell  
;; STREET: 200 E. Broward Boulevard  
;; CITY: Fort Lauderdale  
;; STATE: Florida  
;; COUNTRY: USA  
;; ZIP: 33301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/436,771  
;; FILING DATE:  
;; CLASSIFICATION: 536  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/198,973  
;; FILING DATE: 18-FEB-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Manso, Peter J.  
;; REGISTRATION NUMBER: 32,264  
;; REFERENCE/DOCKET NUMBER: CL11363-16  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 305/527/2498  
;; TELEFAX: 305/764/4996  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2928 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 104...2326  
;; US-08-436-771-3

Query Match 5.3%; Score 61.8; DB 2; Length 2928;  
Best Local Similarity 57.8%; Pred. No. 2.1e-06;  
Matches 130; Conservative 0; Mismatches 92; Indels 3; Gaps 1;  
QY 616 gaaactctgcaagccatggacagacacccaatgtgaggataagctgctgagaccccc 675  
Db 769 GCATCTGCTGCTGACCATGGGCTGATGTCATGAGGGGAGAAAGAGGGAAGACTCC 828  
QY 676 gctgcacgtgagctccgagacagggcaggtggagattgtgagcactttctatccctggg 735  
Db 829 CCTGATCTGCGAGTGGAGAGAGCACTTGGGTTGGTGCAGAGGCTTCTGGAGCAAGA 888  
QY 736 cc---tggaataatccagagacaggggaaggatactgcccctgcatgacgtgtgag 792  
Db 889 GCACATAGAGATTAAATGACACAGACAGTGTGATGCAAAACAGACACTGCTGCTGTGTTGA 948  
QY 793 gctcaacgcgtacaaatcaaatcactgctcctcgtcatggggc 837  
Db 949 ACTCAAACTGAAGAAATGCCGAGTTGCTGTGCAAAACGTGGAGC 993

RESULT 12  
US-08-436-998-1  
; Sequence 1, Application US/08434998  
; Patent No. 5866781  
; GENERAL INFORMATION:  
; APPLICANT: Silverman, Robert H.  
; APPLICANT: Sengupta, Dibyendu N.  
; TITLE OF INVENTION: Antiviral Transgenic Plants, Vectors,  
; TITLE OF INVENTION: Cells and Methods  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:



COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,797  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-16(C)  
TELEPHONE: 305/468-7811  
TELEFAX: 305/463-2030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2928 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 104..2326  
US-08-487-797-3

Query Match 5.3%; Score 61.8; DB 2; Length 2928;  
Best Local Similarity 57.8%; Pred. No. 2.1e-06;  
Matches 130; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

QY 616 gaaacttctgcaaaagccatggagcagacacccaatgtgaggataagctgtgagcacc 675  
DB 769 GCATCTGCTGCTGGACCATGGGCTGATCTCAATGTGAGGGGAGAGGAGGACTCC 828  
QY 676 gctgcacgtggcagtcceggacagggcaggtggagattgtgagcaactttctatccctggg 735  
DB 829 CCTGATCCTGGCAGTGGAGAGAACGACTTGGGTTTGGTGACAGAGCTTCTGGAGCAAGA 888  
QY 736 cc---tgaaatcaatgccagagacagaggaagggataactgccctgcacgtgtgag 792  
DB 889 GCACATAGAGATTAAATGACACAGACAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 948  
QY 793 gctcaaccgctacaaatcaataactgctcctgcacgtgtgagggc 837  
DB 949 ACTCAAACTGAAGAAAATCGCCGAGTGTGCTGTGCAAACTGGAGC 993

Search completed: April 17, 2002, 09:28:45  
Job time: 4769 sec

COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,797  
FILING DATE: 07-JUNE-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: CL11363-16(C)  
TELEPHONE: 305/468-7811  
TELEFAX: 305/463-2030  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2928 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 104..2326  
US-08-487-797-1

Query Match 5.3%; Score 61.8; DB 2; Length 2928;  
Best Local Similarity 57.8%; Pred. No. 2.1e-06;  
Matches 130; Conservative 0; Mismatches 92; Indels 3; Gaps 1;

QY 616 gaaacttctgcaaaagccatggagcagacacccaatgtgaggataagctgtgagcacc 675  
DB 769 GCATCTGCTGCTGGACCATGGGCTGATCTCAATGTGAGGGGAGAGGAGGACTCC 828  
QY 676 gctgcacgtggcagtcceggacagggcaggtggagattgtgagcaactttctatccctggg 735  
DB 829 CCTGATCCTGGCAGTGGAGAGAACGACTTGGGTTTGGTGACAGAGCTTCTGGAGCAAGA 888  
QY 736 cc---tgaaatcaatgccagagacagaggaagggataactgccctgcacgtgtgag 792  
DB 889 GCACATAGAGATTAAATGACACAGACAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 948  
QY 793 gctcaaccgctacaaatcaataactgctcctgcacgtgtgagggc 837  
DB 949 ACTCAAACTGAAGAAAATCGCCGAGTGTGCTGTGCAAACTGGAGC 993

RESULT 15  
US-08-487-797-3  
Sequence 3, Application US/08487797  
Patent No. 5866787  
GENERAL INFORMATION:  
APPLICANT: Silverman, Robert H.  
APPLICANT: Sengupta, Dibyendu N.  
TITLE OF INVENTION: Transgenic Plants Co-Expressing A  
TITLE OF INVENTION: Functional Human 2-5A System  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: One E. Broward Boulevard, #1300  
CITY: Fort Lauderdale  
STATE: Florida  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

---

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 08:09:16 ; Search time 175.48 Seconds  
(without alignments)  
5657.526 Million cell updates/sec

Title: US-09-758-593A-2  
Perfect score: 1158  
Sequence: 1 cagctcgaggagcgaccca.....ctaccacataaaaaagctg 1158

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

arched: 930621 seqs, 428662619 residues  
Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1148.4	99.2	1208	22 AAI59509	Human polynucleoti
2	545.4	47.1	547	22 AAI61295	Human polynucleoti
3	248.4	21.5	1901	22 AAH02910	Human shear stress
4	192.8	16.6	221	21 AAC16163	Human secreted pro
5	113.4	9.8	386	22 AAF66063	Novel human polynu
6	103.8	9.0	363	22 AAF66048	Novel human polynu
7	96.8	8.4	1783	22 AAI60004	Human polynucleoti
8	96.8	8.4	2894	22 AAI58218	Human polynucleoti
9	86.6	7.5	7185	22 AAH57380	Human skeletal mus
10	84.6	7.3	3040	22 AAH76497	Human ORFX ORF2052
11	79.2	6.8	2348	22 AAH46914	cDNA encoding huma

12	77.2	6.7	1288	21	AAZ49052	Human ankryrin fami
13	77.2	6.7	1345	21	AAZ94868	Human MHC class II
14	77.2	6.7	2649	19	AAV38678	Mus musculus SOCS1
15	75.6	6.5	3454	22	AAF83200	Human kidney ankyr
16	74	6.4	33654	22	AAF70259	Human dopamine rec
17	73.8	6.4	869	21	AAC76382	Human ORFX ORF1937
18	73.4	6.3	1959	22	AAH15018	Human cDNA sequenc
19	73.4	6.3	2790	21	AAZ36832	cDNA encoding a su
20	69.8	6.0	5482	22	AAF63955	Drosophila tankyria
21	68.2	5.9	1949	21	AAC77520	Human ORFX ORF3075
22	68.2	5.9	9063	22	AAH23688	Human tumour suppr
23	67	5.8	2499	22	AAS06739	Polynucleotide seq
24	67	5.8	3876	22	AAH15762	Human cDNA sequenc
25	65.8	5.7	1359	22	AAF90197	Nucleotide sequenc
26	65.6	5.7	1160	21	AAZ49069	Ankyrin repeat pro
27	65	5.6	6156	22	AAS01769	Drosophila melanog
28	65	5.6	24358	22	AAS01768	Drosophila melanog
29	64.8	5.6	1797	22	AAF90586	Degenerate DNA cod
30	64.8	5.6	1957	22	AAF90585	cDNA encoding huma
31	64.2	5.5	1780	20	AAZ33623	Human breast tumou
32	64.2	5.5	1786	18	AAT71716	Human integrin-lin
33	64.2	5.5	1789	22	AAF69274	Human integrin-lin
34	63	5.4	2157	18	AAT46188	Bovine Ribonucleas
35	62.8	5.4	1883	22	AAI59595	Human polynucleoti
36	62.4	5.4	16442	18	AAH83006	Partial mouse WRN
37	62.2	5.4	441	21	AAC69723	Human breast tumou
38	62.2	5.4	1773	22	AAH04776	Human death domain
39	62.2	5.4	2429	22	AAH35009	Human colon cancer
40	62.2	5.4	8091	19	AAV57001	Human Notch3 cDNA
41	62	5.4	3984	22	AAF63838	Human tankyrase I
42	62	5.4	4134	21	AAC66826	Human tankyrase I
43	62	5.4	4134	21	AAZ29627	Human tankyrase CD
44	62	5.4	4491	21	AAZ29632	Human truncated ta
45	61.8	5.3	2530	22	AAH72871	Human cervical can

ALIGNMENTS

RESULT 1  
AAI59509  
ID AAI59509 standard; cDNA; 1208 BP.  
XX  
AC AAI59509;  
XX  
DT 22-OCT-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 1712.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
peripheral nervous system; neuropathy; central nervous system; CNS;  
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.

XX Homo sapiens.  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.  
XX 29-NOV-2000; 2000US-0727344.  
XX

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;  
Zhao QA, Zhou P, Goodrich R, Drmanac RT;

WPI: 2001-442253/47.  
P-PSDB; RAM40353.

Novel nucleic acids and polypeptides, useful for treating disorders  
such as central nervous system injuries -

Claim 1; SEQ ID NO 1712; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and  
the encoded polypeptides (RAM38642-RAM42213) with nootropic,  
immunosuppressant and cytoskeletal activity. The polynucleotides are useful  
in gene therapy. A composition containing a polypeptide or polynucleotide  
of the invention may be used to treat diseases of the peripheral nervous  
system, such as peripheral nervous injuries, peripheral neuropathy and  
localised neuropathies and central nervous system diseases, such as  
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
utilisation of the activities such as: immune system suppression,  
Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
assays for receptor activity, arthritis and inflammation, leukaemias and  
C.N.S disorders  
C Note: The sequence data for this patent did not form part of the printed  
specification.

Sequence 1208 BP; 292 A; 335 C; 392 G; 189 T; 0 other;

Query Match 99.2%; Score 1148.4; DB 22; Length 1208;  
Best Local Similarity 99.9%; Pred. No. 1.1e-267;  
Matches 1149; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

9 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 68  
39 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 98  
69 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 128  
99 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 158  
129 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 188  
159 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 218  
189 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 248  
219 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 278  
249 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 308  
279 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 338  
309 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 368  
339 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 398  
369 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 428  
399 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 458  
429 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 488  
459 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 518  
489 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 548  
519 gggagggcaccatgaggaactccgagggcggtgacagagggccacagcgtccatcgagcagc 578

549 attccaggatcgctggaactgcagccatgcatggtggcctgcgcgagggggccacttag 608  
579 attccaggatcgctggaactgcagccatgcatggtggcctgcgcgagggggccacttag 638  
609 aggtggtaaaactctgcacaaagccatggagcagacacacaaatgtgagggataaagctgctga 668  
639 aggtggtaaaactctgcacaaagccatggagcagacacacaaatgtgagggataaagctgctga 698  
669 gaccccgctcagctggcagtcgagcagcagcagcagcagcagcagcagcagcagcagcagc 728  
699 gaccccgctcagctggcagtcgagcagcagcagcagcagcagcagcagcagcagcagcagc 758  
729 ccttgggctggaatcaatccagagacagaggaagggatactgcccctgcatgacgctg 788  
759 ccttgggctggaatcaatccagagacagaggaagggatactgcccctgcatgacgctg 818  
789 tgaggctcaacgcgtacacaaatcaataaactgctgctcctcatgaggggctgacatgatga 848  
819 tgaggctcaacgcgtacacaaatcaataaactgctgctcctcatgaggggctgacatgatga 878  
849 ccaagaacctgaggaagaaagaccccgacgacccctgctgagcagcagcagcagcagcagc 908  
879 ccaagaacctgaggaagaaagaccccgacgacccctgctgagcagcagcagcagcagcagc 938  
909 ggcacgctcaggaactcctgagcgggggctgagcagcagcagcagcagcagcagcagcagc 968  
939 ggcacgctcaggaactcctgagcgggggctgagcagcagcagcagcagcagcagcagcagc 998  
969 atagtggcgagagacccctcagcctgagcagcagcagcagcagcagcagcagcagcagcagc 1028  
999 atagtggcgagagacccctcagcctgagcagcagcagcagcagcagcagcagcagcagcagc 1058  
1029 cagctacccagccctctctgtgtgacagcagcagcagcagcagcagcagcagcagcagcagc 1088  
1059 cagctacccagccctctctgtgtgacagcagcagcagcagcagcagcagcagcagcagcagc 1118  
1089 actgagggccagcagcctttttctgcatgacatccagagagacacataccacaaactaccacaa 1148  
1119 actgagggccagcagcctttttctgcatgacatccagagagacacataccacaaactaccacaa 1178  
1149 aaaaaagctg 1158  
1179 aaaaaagctg 1188  
RESULT 2  
AA161295  
ID AA161295 standard; cDNA; 547 BP.  
AC AA161295;  
DT 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 5284.  
DE Human; nototropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia; ss.  
XX Homo sapiens.  
OS  
XX WO200153312-A1.  
PN  
XX 26-JUL-2001.  
PD  
XX 26-DEC-2000; 2000WO-US34263.  
PF  
XX 21-JAN-2000; 2000US-0488725.  
PR  
XX 25-APR-2000; 2000US-0552317.  
PR



09-JUL-2000; 2000US-0598042.  
 19-JUL-2000; 2000US-0620312.  
 03-AUG-2000; 2000US-0653450.  
 14-SEP-2000; 2000US-0662191.  
 19-OCT-2000; 2000US-0693036.  
 29-NOV-2000; 2000US-0727344.  
 (HYSE-) HYSEQ INC.  
 Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 WPI; 2001-442253/47.  
 P-PSDB; AAM42139.  
 Novel nucleic acids and polypeptides, useful for treating disorders  
 such as central nervous system injuries -  
 Claim 1; SEQ ID NO 5284; 10078pp; English.  
 The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 immunosuppressant and cytostatic activity. The polynucleotides are useful  
 in gene therapy. A composition containing a polypeptide or polynucleotide  
 of the invention may be used to treat diseases of the peripheral nervous  
 system, such as peripheral nervous injuries, peripheral neuropathy and  
 localised neuropathies and central nervous system diseases, such as  
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 utilisation of the activities such as: immune system suppression,  
 Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 assays for receptor activity, arthritis and inflammation, leukaemias and  
 C.N.S disorders.  
 Note: The sequence data for this patent did not form part of the printed  
 specification.  
 Sequence 547 BP; 132 A; 145 C; 195 G; 75 T; 0 other;

Query Match 47.1%; Score 545.4; DB 22; Length 547;  
 Best Local Similarity 99.8%; Pred. No. 2.3e-122;  
 Matches 546; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 11 gacggcaccatgagactccgagcggtgacagggccacacgagcgtcatcgagcgg 70  
 Db 1 gacggcaccatgagactccgagcggtgacagggccacacgagcgtcatcgagcgg 60  
 Qy 71 ctggccacaggaggaggaatgagaaactccgagagagacacacgagcgtgccatg 130  
 Db 61 ctggccacaggaggaggaatgagaaactccgagagagacacacgagcgtgccatg 120  
 Qy 131 gacttgctggtctgagagatgagagacacacgaggggtcagagtgccctgcagaag 190  
 Db 121 gacttgctggtgctgagagatgagagacacacgaggggtcagagtgccctgcagaag 180  
 Qy 191 gtgaaggcccaagagcggtgacagacgctccctggacctgcggtgagatcatcgat 250  
 Db 181 gtgaaggcccaagagcggtgacagacgctccctggacctgcggtgagatcatcgat 240  
 Qy 251 gtggcgggatccagaacctatcagctgcggaagaaacgcaagcagaagaagcgggac 310  
 Db 241 gtggcgggatccagaacctatcagctgcggaagaaacgcaagcagaagaagcgggac 300  
 Qy 311 gctctggccgctcagatgagccgccccagagcccgagagatcactggccctgtgat 370  
 Db 301 gctctggccgctcagatgagccgccccagagcccgagagatcactggccctgtgat 360  
 Qy 371 gaggagacctctctgaaagctcggtgaggggaaatgaaagtcattgagaagttccgt 430  
 Db 361 gaggagacctctctgaaagctcggtgaggggaaatgaaagtcattgagaagttccgt 420

Qy 431 gctgacgggggtcagcgacacacgtgacaccagcttccgtcgcagacagcactgaccagact 490  
 Db 421 gctgacgggggtcagcgacacacgtgacaccagcttccgtcgcagacagcactgaccagact 480  
 Qy 491 tccctggaaggccacatggaatcctctggaagcgttctagataatggggccactgtggac 550  
 Db 481 tccctggaaggccacatggaatcctctggaagcgttctagataatggggccactgtggac 540  
 Qy 551 ttccagg 557  
 Db 541 ttccagg 547  
 RESULT 3  
 AAH02910  
 ID AAH02910 standard; DNA; 1901 BP.  
 XX  
 AC AAH02910;  
 XX  
 DT 15-JUN-2001 (first entry)  
 XX  
 DE Human shear stress-response coding sequence SEQ ID NO: 73.  
 XX  
 KW Human; shear stress-response protein; vascular disease;  
 KW arteriosclerosis; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200125427-A1.  
 XX  
 PD 12-APR-2001.  
 XX  
 PF 02-OCT-2000; 2000WO-JP06840.  
 XX  
 PR 01-OCT-1999; 95JP-0280976.  
 XX  
 PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 PA (NOJI/) NOJIMA H.  
 PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;  
 PI Kuga T, Sekine S, Nakamura Y, Sugano S;  
 XX  
 DR WPI; 2001-266308/27.  
 DR P-PSDB; AAB90787.  
 XX  
 PT DNA sequences, proteins encoded by them and antibodies against them  
 PT useful in diagnosis and treatment of vascular disease caused by  
 PT arteriosclerosis -  
 XX  
 PS Claim 20; Page 422-425; 678pp; Japanese.  
 XX  
 CC The present invention provides the protein and coding sequences of a  
 CC number of human shear stress response proteins. These are useful in the  
 CC diagnosis, treatment and screening of vascular diseases caused by  
 CC arteriosclerosis, including heart failure, post-PTCA restenosis and  
 CC hypertension.  
 XX  
 SQ Sequence 1901 BP; 592 A; 378 C; 460 G; 471 T; 0 other;  
 Query Match 21.5%; Score 248.4; DB 22; Length 1901;  
 Best Local Similarity 64.7%; Pred. No. 1.4e-50;  
 Matches 369; Conservative 0; Mismatches 201; Indels 0; Gaps 0;  
 Qy 338 ccagagcccgagagatcactggcctgtggatgagagacaccttctctgaagctgcgtg 397  
 Db 577 ccagagcccgagagatcactggcctgtggatgagagacaccttctctgaagctgcgtg 636  
 Qy 398 gaggggaaaatgaagtcattgagaagttctctggctgacgggggtcagccgacacgtgc 457  
 Db 637 gagaataaactccagtagtagaaaaattctgtcagacaaagaacaatccagatgttgt 696  
 Qy 458 gaccagttccctgcgacagcactgaccagcttccctggaaggccacatggaatccctg 517

Db 697 gatgataaaacgacagctctctatagacatgcttgaagacatttggaattgtg 756  
 QY 518 gagaagcttctagataatggggccactgtggacttccagatcgctgactgcacagcc 577  
 Db 757 gagaagttaataagactggagcccgatcgaaattccgtgatgtgttaatccacagcc 816  
 QY 578 atcattggcctccgctggggccacttagagtggtgaaacttctgcaaacccatgga 637  
 Db 817 atccactggcgaagcgtggaggaacccggatgttttaaaattgttctgaataaagg 876  
 QY 638 gcagacaccaatgtgaggataagctgtgagcaccgccctgcacgtgcagtcgcgaca 697  
 Db 877 gcaaaatagcccgagataaattgtctcagcacagcgtgctggtgagact 936  
 QY 698 gggcagggtgagattgtgagcacttctatccctggcctggaataatccaagac 757  
 Db 937 ggcactatgagtcgagcagactctatccctgtgagcagacctcaagcccaagac 996  
 QY 758 aggaaggggatactccctgcactgacgtgtgagctcaacccgtacaaaatcatcaa 817  
 Db 997 agagaaggagatacccccgttgcattgacgtgtgagactgaaccgtataagatgccga 1056  
 QY 818 ctgactctctgactgagcgtgacatgatgacaaagaactggcagaaagaccccgacg 877  
 Db 1057 ctctgattatgctgagcggcgatctcaacatcaagaactgtgctgggaagacgcgatg 1116  
 QY 878 gacctggtgcagctctggcagcgtgatacc 907  
 Db 1117 gatctggtgtacactggcagaatgaacc 1146

RESULT 4

AAC16163  
 ID AAC16163 standard; cDNA; 221 BP.  
 XX AAC16163;  
 AC AAC16163;  
 DT 06-OCT-2000 (first entry)  
 XX Human secreted protein 5' EST, SEQ ID NO: 20238.  
 DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN EP1033401-A2.  
 XX 06-SEP-2000.  
 PF 21-FEB-2000; 2000EP-0200610.  
 PR 26-FEB-1999; 99US-0122487.  
 PA (GEST) GENSET.  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI WPI; 2000-500381/45.  
 DR New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 XX obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX Claim 1; SEQ ID 20238; 71pp + CD-ROM; English.  
 XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA

CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.  
 XX Sequence 221 BP; 54 A; 54 C; 87 G; 24 T; 2 other;

Query Match 16.6%; Score 192.8; DB 21; Length 221;  
 Best Local Similarity 98.6%; Pred. No. 1.9e-37;  
 Matches 205; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 8 agggacggccaccatggaggactcccgagcggtgcagaggccacagcgtcatcgagcag 67  
 Db 11 atggacggccaccatggaggactcccgagcggtgcagaggccacagcgtcatcgagcag 70  
 QY 68 cgcctggcacagagagagaaatgagaaactccgaggagacacacg-ccagaagctgcc 126  
 Db 71 cggctggcacagagagagaaatgagaaactccgaggagacacg-ccagaagctgcc 130  
 QY 127 catgacttgcgtgctggaggtgagaaagcaccacgggctcagagtgagcagccctgca 186  
 Db 131 catgacttgcgtgctggaggtgagaaagcaccacgggctcagagtgagcagccctgca 190  
 QY 187 gaagtgaaaggccaaagagcgcgtgcgc 214  
 Db 191 gaagtgaaaggccaaagagcgcgtgcgc 218

RESULT 5

AAF66063  
 ID AAF66063 standard; cDNA; 386 BP.  
 XX AAF66063;  
 AC AAF66063;  
 DT 09-APR-2001 (first entry)  
 XX Novel human polynucleotide, SEQ ID NO: 1819.  
 DE Human; cytostatic; gene therapy; colon cancer; prostate cancer;  
 KW breast cancer; lung cancer; cancer detection; ss.  
 KW Homo sapiens.  
 OS Homo sapiens.  
 PN WO200102568-A2.  
 XX 11-JAN-2001.  
 PF 30-JUN-2000; 2000WO-US18374.  
 XX 02-JUL-1999; 99US-0142310.  
 PR 02-JUL-1999; 99US-0142311.  
 XX (CHIR) CHIRON CORP.  
 PA (HYSE-) HYSEQ INC.  
 XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;  
 PI Reihard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;  
 PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;  
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;  
 XX WPI; 2001-091805/10.  
 DR Library of polynucleotides for diagnosing a cancerous state of a  
 XX mammalian cell and detecting cancer, particularly of the colon or  
 PT prostate, comprises 3351 human polynucleotide sequences -  
 XX Claim 9; Page 804; 1046pp; English.

CC The present sequence is one of 3351 sequences in a library of human  
 CC polynucleotides. The library is used to detect differentially expressed  
 CC genes correlated with a cancerous state of a mammalian cell and can  
 CC detect colon, prostate, breast and lung cancer. The library can be used  
 CC to produce probes for detection of mRNA and to produce additional copies  
 CC of the polynucleotides. The probes can be used for chromosome mapping of  
 CC the polynucleotide and for detection of transcription levels. Ribozymes  
 CC or antisense oligonucleotides can be generated. The polynucleotides and  
 CC their gene products are used as genetic or biochemical markers (e.g. in  
 CC blood or tissues) that will detect the earliest changes along the  
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and  
 CC preventive interventions. The polynucleotides, polypeptides and  
 CC antibodies against them can be used in pharmaceutical compositions to  
 CC treat the cancers and proliferative disorders such as neoplasia,  
 CC dysplasia and hyperplasia.  
 XX  
 SQ Sequence 386 BP; 135 A; 68 C; 97 G; 86 T; 0 other;

Query Match 9.8%; Score 113.4; DB 22; Length 386;  
 Best Local Similarity 63.3%; Pred. NO. 3.1e-18;  
 Matches 174; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
 QY 338 ccagagccggaggagatacctggccctgtgatgagggagaccttctctgaaagctgcggtg 397  
 DB 112 ccagaaacctgaatacattacggaacctgtggtgctacgtttctgaaagctgctgtg 171  
 QY 398 gaggggaaaaatgaaggtcattgagagttcctgctgacgggggtcagcgacacgtgc 457  
 DB 172 gagaataaactgcagtagtagaataattctgtcagacaagaacaatccagatgttctg 231  
 QY 458 gaccagttccctgcggacagcactgcaccgagcttccctggaaggccacatggaatcctg 517  
 DB 232 gatgagtataaacggacagctcttcctatagagcagtctgtggaaggacattggcaattgtg 291  
 QY 518 gagaagcttctagataaaggggccactgtggacttccaggtatcggtggaactgcacagcc 577  
 DB 292 gagaagtttaagtgaagctggagccagatcgaattccgtgatgtatgttgaatccacagcc 351  
 QY 578 atgcattggcctgcgcggggggccacttagaggt 612  
 DB 352 atccactgggcaagcctggaggaaaccttgatg 386

RESULT 6  
 AAF66048  
 ID AAF66048 standard; cDNA; 363 BP.  
 XX  
 AC AAF66048;  
 XX  
 DT 09-APR-2001 (first entry)  
 XX  
 DE Novel human polynucleotide, SEQ ID NO: 1804.  
 XX  
 KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;  
 KW breast cancer; lung cancer; cancer detection; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200102568-A2.  
 XX  
 PD 11-JAN-2001.  
 XX  
 PF 30-JUN-2000; 2000WO-US18374.  
 XX  
 PR 02-JUL-1999; 99US-0142310.  
 PR 02-JUL-1999; 99US-0142311.  
 XX  
 PA (CHIR) CHIRON CORP.  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;  
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;

PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;  
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;  
 XX  
 DR WPT; 2001-091805/10.  
 XX  
 PT Library of polynucleotides for diagnosing a cancerous state of a  
 PT mammalian cell and detecting cancer, particularly of the colon or  
 PT prostate, comprises 3351 human polynucleotide sequences -  
 XX  
 PS Claim 9; Page 802; 1046pp; English.  
 XX  
 CC The present sequence is one of 3351 sequences in a library of human  
 CC polynucleotides. The library is used to detect differentially expressed  
 CC genes correlated with a cancerous state of a mammalian cell and can  
 CC detect colon, prostate, breast and lung cancer. The library can be used  
 CC to produce probes for detection of mRNA and to produce additional copies  
 CC of the polynucleotides. The probes can be used for chromosome mapping of  
 CC the polynucleotide and for detection of transcription levels. Ribozymes  
 CC or antisense oligonucleotides can be generated. The polynucleotides and  
 CC their gene products are used as genetic or biochemical markers (e.g. in  
 CC blood or tissues) that will detect the earliest changes along the  
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and  
 CC preventive interventions. The polynucleotides, polypeptides and  
 CC antibodies against them can be used in pharmaceutical compositions to  
 CC treat the cancers and proliferative disorders such as neoplasia,  
 CC dysplasia and hyperplasia.  
 XX  
 SQ Sequence 363 BP; 128 A; 63 C; 87 G; 85 T; 0 other;

Query Match 9.0%; Score 103.8; DB 22; Length 363;  
 Best Local Similarity 63.3%; Pred. NO. 6.3e-16;  
 Matches 159; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
 QY 338 ccagagccggaggagatacctggccctgtgatgagggagaccttctgaaagctgcggtg 397  
 DB 112 ccagaaacctgaatacattacggaacctgtggtgctacgtttctgaaagctgctgtg 171  
 QY 398 gaggggaaaaatgaaggtcattgagagttcctgctgacgggggtcagcgacacgtgc 457  
 DB 172 gagaataaactgcagtagtagaataattctgtcagacaagaacaatccagatgttctg 231  
 QY 458 gaccagttccctgcggacagcactgcaccgagcttccctggaaggccacatggaatcctg 517  
 DB 232 gatgagtataaacggacagctcttcctatagagcagtctgtggaaggacattggcaattgtg 291  
 QY 518 gagaagcttctagataaaggggccactgtggacttccaggtatcggtggaactgcacagcc 577  
 DB 292 gagaagtttaagtgaagctggagccagatcgaattccgtgatgtatgttgaatccacagcc 351  
 QY 578 atgcattgggc 588  
 DB 352 atccactgggc 362

RESULT 7  
 AAI60004  
 ID AAI60004 standard; cDNA; 1783 BP.  
 XX  
 AC AAI60004;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 3993.  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.

XX WO200153312-A1.  
 PN 26-JUL-2001.  
 PD  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 XX WPI: 2001-442253/47.  
 DR P-PSDB; AAM40848.  
 DR  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Claim 1; SEQ ID NO 3993; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 XX Sequence 1783 BP; 508 A; 383 C; 446 G; 446 T; 0 other;

Query Match 8.4%; Score 96.8; DB 22; Length 1783;  
 Best Local Similarity 55.7%; Pred. No. 4.9e-14;  
 Matches 185; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 455 tgcgaccagttccgttcgacagcactgcaccgagcttcctcctggaagccacatggaaatc 514  
 DB 282 tcgcaccagggggcgccagccttgccaccatggcctctgaacggccacgtggagatg 341  
 QY 515 ctggagaagcttcagataatggggccactgtggacttcacagatggctggactgcaca 574  
 DB 342 gtcaattactctggccaaaggggcaaatatacaatgcatttgacaagaagcaggcggt 401  
 QY 575 gccatgcattggcctgcgcggggccacttagaggttggtgaaactctgcagaagccat 634  
 DB 402 gctctgcactgggcagcaccatgcacacttggtgattgtgtagcattgtcattaacct 461  
 QY 635 gagcagacacacaaatgtgaggataaagtgtgcagacacccctgcacgtggcagtcagg 694  
 DB 462 ggccgagaagtgacctgtgaagataaagggttataccctctctgcagtcgacctcc 521  
 QY 695 acaggcaggtgagattgtgagcacttctatccctggcctggcctggaatcaatgcaga 754  
 DB 522 aatggacagattaatgtttcagaagatctcctgaacctgggggtggagattgatgaatc 581

QY 755 gacagggaaggagatactgacctgcctgcagtcgcgc 786  
 DB 582 aatgtctatggaataacacagcgttcacatcgc 613

RESULT 8  
 AAI58218  
 ID AAI58218 standard; cDNA; 2894 BP.  
 XX  
 AC AAI58218;  
 XX  
 DT 22-OCT-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 421.  
 XX  
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200153312-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 26-DEC-2000; 2000WO-US34263.  
 XX  
 PR 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 09-JUL-2000; 2000US-0598042.  
 PR 19-JUL-2000; 2000US-0620312.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX  
 XX WPI: 2001-442253/47.  
 DR P-PSDB; AAM39062.  
 DR  
 XX  
 PT Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX  
 PS Claim 1; SEQ ID NO 421; 10078pp; English.  
 XX  
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX  
 XX Sequence 2894 BP; 808 A; 627 C; 741 G; 718 T; 0 other;





PT Kinase polypeptides useful for treating cancers, Alzheimer's disease,  
PT viral infections, diabetes, obesity, organ transplant rejection and  
XX rheumatoid arthritis -  
PS Example 1; Page 210; 218pp; English.  
XX  
CC The invention provides human protein kinases and protein kinase-like  
CC enzymes and polynucleotides encoding the polypeptides. The kinase  
CC polypeptides and their modulators are useful for treating a disease or  
CC disorder such as cancer, immune-related diseases, cardiovascular disease,  
CC brain or neuronal-associated disease and metabolic disorders, including  
CC cancers of tissues, cancers of hematopoietic origin, diseases of the  
CC central nervous system, diseases of the peripheral nervous system,  
CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic  
CC lateral sclerosis, viral infections, infections caused by prions,  
CC bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,  
CC mood disorders, attention disorders, cognitive disorders, hypotension,  
CC hypertension, psychotic disorders, neurological disorders, dyskinesias,  
CC metabolic disorders, and organ transplant rejection. They are also useful  
CC for treating rhinitis, autoimmunity, atherosclerosis, psoriasis,  
CC osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic  
CC inflammatory bowel disease, rheumatoid arthritis, metabolic disorders  
CC such as diabetes, obesity, cardiovascular diseases such as reperfusion  
CC injury, coronary thrombosis, clotting disorders and atherosclerosis,  
CC ocular diseases such as glaucoma, retinopathy and macular degeneration,  
CC psychiatric and neurological disorders such as anxiety, schizophrenia,  
CC dementia, manic depression, etc. The polynucleotides are useful in gene  
CC therapy techniques to treat the above mentioned disorders. Sequences  
CC AAH46891-46922 represent human protein kinases encoding cDNA molecules.  
XX  
SQ Sequence 2348 BP; 516 A; 717 C; 696 G; 419 T; 0 other;

Query Match 6.8%; Score 79.2; DB 22; Length 2348;  
Best Local Similarity 47.2%; Pred. No. 9.3e-10;  
Matches 240; Conservative 0; Mismatches 268; Indels 0; Gaps 0;  
Qy 389 gctcgggtgaggggaaatgaagtcattggaagttcctgctgacgaggggtcagcc 448  
Db 1752 gcaagtgcaggggcaaatatcgtatcgtcaagatgctgctcaggtaaggcagcctt 1811  
Qy 449 gacagctgcagcagcttcctgcagacagcactgcacagagcttccctggaaggccactg 508  
Db 1812 gactgctcccccaggggctgcagacccctgcatctgacagctcacaaggccactg 1871  
Qy 509 gaactcctggagagcttctagataatgggcccactgtggacttcacagatcggtctgac 568  
Db 1872 gagatcatcctctgctgcagagagcagcacaacatgggtgctcttggagctgtgac 1931  
Qy 569 tgcacagccatgcattgggctgcgcggggggccacttagaggtggtgaaactcttgcac 628  
Db 1932 tggactccctgcactagctgcacgcacggggagggcggtggtgtcagcactgctg 1991  
Qy 629 agccatgagcagacacaaatgtaggataagctgctgacacccctcactgactgca 688  
Db 1992 cagtgtgggctgaccccaatgctgcagacagtcagctggacacccctccactctgcy 2051  
Qy 689 gtcgcagagggcaggtgagattgtgagcacttctatccctgggctggaaatcaat 748  
Db 2052 gtccagagagcacttctgagtgcatcactcctcctagacatcagcaaatgtccac 2111  
Qy 749 gccagagagagggggaactgctcctgcagtcagctgtgaggtcacaocgtcacaaa 808  
Db 2112 gccgcacaaggtggtggtgacacccgccacccctgcccgcctcaagggaacacagcc 2171  
Qy 809 atcatcaactgctcctgctgaggtgctgacatgacagacacacctggcaggaag 868  
Db 2172 atctcacaagtgtgtgagggcagcgccagctggagctccaggtgaggtgagctgc 2231  
Qy 869 accccgcagcactggtgcagctctggc 896  
Db 2232 acacccctgcaactggcctccgcagcc 2259

RESULT 12  
AAZ49052  
ID AAZ49052 standard; DNA; 1288 BP.  
XX  
AC AAZ49052;  
XX  
DT 03-APR-2000 (first entry)  
XX  
DE Human ankyrin family protein, ANFP, coding sequence.  
XX  
KW Human; ankyrin family protein: ANFP; autoimmune disorder; inflammation;  
KW atherosclerosis; inflammatory disorder; proliferative disorder; AIDS;  
KW vesicle-trafficking disorder; allergy; amyloidosis; anaemia; asthma;  
KW bronchitis; Crohn's disease; atopic dermatitis; diabetes mellitus;  
KW irritable bowel syndrome; osteoporosis; rheumatoid arthritis; cirrhosis;  
KW hepatitis; ulcerative colitis; cancer; hypercholesterolemia; therapy;  
KW diagnosis; ss.  
XX  
OS Homo sapiens.  
XX  
PN US989863-A.  
XX  
PD 23-NOV-1999.  
XX  
PF 14-OCT-1998; 98US-0172977.  
XX  
PR 14-OCT-1998; 98US-0172977.  
XX  
PA (INCY-) INCYTE PHARM INC.  
XX  
PI Tang YT, Corley NC, Yue H, Guegler KJ;  
XX  
DR WPI: 2000-095634/08.  
DR P-PSDB: AAY59539.  
XX  
PT Polynucleotide sequence encoding a human ankyrin family protein useful  
PT for diagnosis or treatment of autoimmune, inflammatory, proliferative,  
PT and vesicle-trafficking disorders  
XX  
PS Claim 5; Fig 1; 34pp; English.  
XX  
CC This sequence encodes the human ankyrin family protein, ANFP, of the  
CC invention. Host cells containing an expression vector containing the  
CC polynucleotide sequence can be cultured to produce ANFP, which can be  
CC used for diagnosis or treatment of autoimmune, inflammatory,  
CC proliferative and vesicle-trafficking disorders. Disorders which can be  
CC treated include acquired immune deficiency syndrome (AIDS), allergies,  
CC amyloidosis, anaemia, asthma, atherosclerosis, bronchitis, Crohn's  
CC disease, atopic dermatitis, diabetes mellitus, irritable bowel syndrome,  
CC myocardial or pericardial inflammation, osteoporosis, rheumatoid  
CC arthritis, cirrhosis, hepatitis, ulcerative colitis, cancer and  
CC hypercholesterolemia. The polynucleotide sequences can also be used as a  
CC hybridisation probe to detect ANFP-encoding polynucleotides in biological  
CC samples. Purified ANFP can be used to produce antibodies or to screen  
CC libraries of pharmaceutical agents to find agents that specifically bind  
CC ANFP. The DNA and its antisense sequence can be used in therapeutic  
CC compositions e.g. to regulate gene function. The DNA sequence can be used  
CC for diagnostic purposes to detect and quantitate gene expression in  
CC biopsied tissues and to indicate the absence, presence and excess  
CC expression of ANFP and monitor its levels during therapeutic  
CC intervention.  
XX  
SQ Sequence 1288 BP; 284 A; 390 C; 391 G; 223 T; 0 other;

Query Match 6.7%; Score 77.2; DB 21; Length 1288;  
Best Local Similarity 48.2%; Pred. No. 2.4e-09;  
Matches 217; Conservative 0; Mismatches 233; Indels 0; Gaps 0;  
Qy 435 acggggggtcagccgacacgctgcgaccaggttcctgcgcagcagcagctgcacagctccc 494  
Db 589 acgaggtgtcagctctgcgcgcccacccctagacctcctgtccatccaccagctcgcagcac 648





Db 656 acgagggtcagctcgcggccaccctagactccctgtccatccaccagctgcagcac 715  
Qy 495 tgaagccacatgaaatcctcgtgagaagctcttagataatggggccactgtgacttcc 554  
Db 716 agggggagctggaccagctgaaggagcatttgcggaaagtgcacactcgtccacaagc 775  
Qy 555 aggatcggtggactgcacagccatgctatgggctgcgcggggccacttagagtg 614  
Db 776 cagacgagcgcgcttcaccccctcatctggtcctccgcttggagagattgagaccg 835  
Qy 615 tgaacttcgaaagccatgagcagacacccaatgtgaggataagctgctgagcacc 674  
Db 836 ttgcctcctcgtggagtgggtggtccaccccatcctggcaaaagagcgagagcg 895  
Qy 675 cgctgcagctggcagtcgacagcgagctggagattgtgagcactttctatccctgg 734  
Db 896 ccctgctgcgctgcacagcgcgctacacagacattgtgggctgctgagagctg 955  
735 gcttggaaatacaatgcagagacagggaagggtatctgacctgcacgtgagggc 794  
Db 956 acgtggacatacaatctatgtatggaatggaggagcgccactgtgtacgctgtgcgcg 1015  
Qy 795 tcaaccgtcaaaaatacaataactgctcctcgtgatgggctgagctgacatgatgaccaaga 854  
Db 1016 ggaaccacgtgaaatgcgttgaggccttgcgtggcccgaggcgctgaacctaccaccgaag 1075  
Qy 855 acctggcaggaaagaccccgacgacctgg 884  
Db 1076 ccgactcgtgctacaccccgatgaccttg 1105

## RESULT 14

AAV38678  
ID AAV38678 standard; DNA; 2649 BP.

XX AC AAV38678;

XX DT 27-OCT-1998 (first entry)

XX DE Mus musculus SOCS10 gene.

XX KW SOCS; suppressor of cytokine signalling; PCR primer;  
XX KW autoimmune disease; diagnosis; cancer; treatment;  
XX KW cytokine mediated cellular responsiveness; hyperimmunity;  
XX KW immunosuppression; allergies; hypertension; ss.

XX DS Mus musculus.

XX PN WO9820023-A1.

XX PD 14-MAY-1998.

XX PF 31-OCT-1997; 97WO-AU00729.

XX PR 14-FEB-1997; 97AU-0005117.

XX PR 01-NOV-1996; 96AU-0003384.

XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.

XX PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE;  
XX PI Nicola NA, Richardson RT, Starf R, Viney EM, Willson TA;

XX DR WPI; 1998-286854/25.

XX PT Suppressor of cytokine signalling proteins - useful to treat  
XX PT disease, injury or abnormality involving cytokine mediated cellular  
XX PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and  
XX PT hypertension

XX PS Claim 14; Page 153-155; 325pp; English.

XX CC The sequence is that of a gene encoding a suppressor of cytokine

CC signalling protein (SOCS). SOCS can be used to screen for naturally  
CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune  
CC diseases. Alternatively, specific antibodies can be used to  
CC screen for SOCS, which is useful as a knowledge of SOCS levels  
CC may be important for the diagnosis of certain cancers. Soluble  
CC SOCS polypeptides can be used to treat disease, injury or  
CC abnormality involving cytokine mediated cellular responsiveness,  
CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.  
XX XX  
SQ Sequence 2649 BP; 559 A; 791 C; 749 G; 550 T; 0 other;

Query Match 6.7%; Score 77.2; DB 19; Length 2649;  
Best Local Similarity 48.4%; Pred. No. 2.9e-09;  
Matches 214; Conservative 0; Mismatches 228; Indels 0; Gaps 0;

Qy 386 aaagctcggtggagggaataaaggtcattgagaagttcctggctgacgggggtca 445  
Db 619 aaagcctgtgagcgcaagaacgcggggcgtgaggatattggcgatatacaacgcagac 678  
Qy 446 gccgacacgtgcgaccagttccgtcggacagcactgcaccgagcttccctggaagggcac 505  
Db 679 gccaacaccgctgtaacagggctggaccgactgcacgagctgtctcccgcaatgac 738  
Qy 506 atgaaatcctcggagaagcttctagataatggggccactgtggaactccagatcgctg 565  
Db 739 ctggaggctcatggagatcctagtgtgctggcgggccaaaggtggagccaagaatgtctac 798  
Qy 566 gactgcacagccatgcttggcctgcgcggggccacttagagtggtgaaacttctg 625  
Db 799 agcatcacccttctgtgtgctgccagagtgggcagctggagggcctgaggtctctg 858  
Qy 626 caaagccatggagcagacaccaatgtgagggataagctgtgagcaccctcgtgcactg 685  
Db 859 gccaaagcatgtgcagacatcaacacgcagccagtgacagtgcacatcagcctctacag 918  
Qy 686 gcagtccggcagggcaggtgagattgtgagcaacttctatccctgggctggaatac 745  
Db 919 gccagcaagaatgagcatgaagacgtggtagatttctctctcagggcgcgatgct 978  
Qy 746 aatgccagagacaggaagggtatactgcctcgtcatgcgtgtgagctcaaccgctac 805  
Db 979 acaaaagccaacaggcgctgctccctcgtcatgttgcctccaagaagggaactat 1038  
Qy 806 aaatcatcaaaactgctgctcc 827  
Db 1039 agaatagtgcagatgctgctgc 1060

## RESULT 15

AAF83200  
ID AAF83200 standard; DNA; 3454 BP.

XX AC AAF83200;

XX DT 09-JUL-2001 (first entry)

XX DE Human kidney ankyrin G119 (AnkG119) protein encoding DNA.

XX KW Ankyrin G119; AnkG119; kidney; human; spectrin binding domain;  
XX KW regulatory domain; integral membrane protein; secretory protein; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX FT CDS 131..3397  
XX FT /\*tag= a  
XX FT /\*product= "AnkG119"

XX PN US6225086-B1.

XX PD 01-MAY-2001.

XX XX

```
PF 21-MAY-1998; 98US-0082059.
XX
PR 21-MAY-1997; 97US-0047356.
XX
PA (UYUA ) UNIV YALE.
XX
PI MORROW JS, Devarajan P;
XX
WPI; 2001-315683/33.
DR P-PSDB; AAB62429.
XX
PT Novel isolated nucleic acid encoding ankyrin G119 protein, useful for
PT producing recombinant ankyrin G119 protein which participates in
PT trafficking of secretory proteins between endoplasmic reticulum and
PT compartments
XX
XX Claim 4; Fig 2A-C; 40pp; English.
XX
CC The invention relates to a human ankyrin G119 (AnkG119) protein. The
CC AnkG119 has a molecular weight of 116 kDa and comprises a region
CC consisting of 13 repeats of a 33 residue structure and further comprises
CC a spectrin binding domain and a 5 kDa regulatory domain. The AnkG119
CC nucleic acid is useful for producing the protein by standard recombinant
CC techniques. The AnkG119 protein participates in the trafficking of
CC integral membrane proteins and secretory proteins between the endoplasmic
CC reticulum and other membrane compartments. The present sequence
CC represents the DNA encoding the human kidney Ankg119 protein.
XX
XX Sequence 3454 BP; 1043 A; 802 C; 797 G; 812 T; 0 other;
XX

Query Match 6.5%; Score 75.6; DB 22; Length 3454;
Best Local Similarity 47.5%; Pred. NO. 7.7e-09;
Matches 225; Conservative 0; Mismatches 249; Indels 0; Gaps 0;

QY 378 cttctgaaagctcggtgaggaggaataaaggtcaattgagaagttcctggtgacg 437
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 993 ccgtccatctcagctcaggaaggagcggtgacatggtcgtcgtctcctcgtagaa 1052
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 438 gggggtcagcgcagctgacacgtgaccagttccctcgacagcactgcacgagcttccctgg 497
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1053 atgcgaatgtgaacctgacgaataagagcggtgacccactcattggtggtcgaag 1112
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 498 aaggccacatgaaatcctcggagagctctctagataatggggccactgtggactccagg 557
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1113 aagatcgagtgaatgtggcagaagtcctcgttaacacaaagggtcctatgtggacgccaga 1172
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 558 atcggtgactgcagcagcattggtggtcctgcgcggggggccacttagaggtgga 617
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1173 caaagatggatacacacacactgcatgtggtggtcctatggaataatcaagattgta 1232
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 618 aacttctgaaagccatggagcagacaccaaattgaggagataagctgctgagcaccgcg 677
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1233 atttctgtccagcatctgcataaagttaatgccaaacaagaatgggtatagccat 1292
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 678 tgcagctggcagtcggcagcagggcaggtgagattgtgagcactttctatccctgggcc 737
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1293 tacataagcagcagcagcaggggcatacgcatataataaatgtcttacttcgaacaacg 1352
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 738 tggaaatacaatccagagacaggaaggagatactgcccctgcatgacgtgtgaggtcca 797
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1353 cctccccaatgaactcactgtgaatgggaatactgccccttggtgattgcccggtgctcg 1412
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 798 accgtacaaaatcatcaaatgctgctcctgcatgggggtgacatgacca 851
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1413 gctacatctcagtagtgacacccctgaagatagtagccgaagaacacatgacca 1466
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: April 17, 2002, 09:31:29  
Job time: 4933 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 17, 2002, 08:09:16 ; Search time 1550.75 Seconds  
(without alignments)  
8024.254 Million cell updates/sec

Title: US-09-758-593A-2

Perfect score: 1158

Sequence: 1 cagctcgaggacgcacca.....ctaccacaataaaaaagctg 1158

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

arched: 11351937 seqs, 5372889281 residues

total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estfun:\*
  - 2: em\_esthum:\*
  - 3: em\_estin:\*
  - 4: em\_estom:\*
  - 5: em\_estpl:\*
  - 6: em\_estba:\*
  - 7: em\_estro:\*
  - 8: em\_estov:\*
  - 9: em\_htc:\*
  - 10: gb\_est1:\*
  - 11: gb\_est2:\*
  - 12: gb\_htc:\*
  - 13: gb\_gss:\*
  - 14: em\_gss\_fun:\*
  - 15: em\_gss\_hum:\*
  - 16: em\_gss\_inv:\*
  - 17: em\_gss\_pln:\*
  - 18: em\_gss\_pro:\*
  - 19: em\_gss\_rod:\*
  - 20: em\_gss\_vrt:\*
  - 21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568.4	49.1	749	11	BF790486 602250156
2	548.6	47.4	610	10	AW044467 wx22c10.x
3	530.2	45.8	875	11	BI413513 602987535
4	492.4	42.5	504	11	BE926721 QV2-BR063
5	433.4	37.4	456	10	AJ346356 shr-00001
6	423.8	36.6	435	11	F24585 HSPD11011 H
7	420.4	36.3	567	11	BF189893 235854 MA
8	415	35.8	417	11	F32675 HSPD25601 H
9	410.2	35.4	457	11	F36391 HSPD33981 H
10	402	34.7	408	10	AJ346328 shr-00001
11	399.2	34.5	423	10	AJ346414 shr-00001
12	393.6	34.0	537	11	BF189892 235852 MA

13	390.4	33.7	504	11	BE809163
14	384.4	33.2	412	10	AJ346604
15	383.2	33.1	489	11	BF775625
16	382	33.0	408	10	AA180314
17	379.4	32.8	385	11	F25326
18	369.4	31.9	380	11	F28193
19	366.2	31.6	371	10	AJ346337
20	354	30.6	372	10	AJ346193 shr-00001
21	352.4	30.4	354	10	AJ346514 shr-00001
22	348	30.1	350	11	F28615
23	348	30.1	360	11	F28533
24	323	27.9	324	11	F24152
25	318	27.5	341	11	F19619
26	314	27.1	326	10	AW237359
27	307.4	26.5	325	10	AJ346541 shr-00001
28	307	26.5	321	10	AJ346516 shr-00001
29	306.8	26.5	391	11	BF599511
30	299	25.8	302	10	AJ346216 shr-00001
31	299	25.8	321	10	AJ345979 shr-00001
32	297	25.6	299	10	AJ346507 shr-00001
33	297	25.6	299	10	AJ346764 shr-00001
34	297	25.6	301	10	AJ346249 shr-00001
35	297	25.6	308	10	AJ346729 shr-00001
36	297	25.6	308	10	AJ346755 shr-00001
37	297	25.6	309	10	AJ346346 shr-00001
38	297	25.6	310	10	AJ346006 shr-00001
39	297	25.6	311	10	AJ346355 shr-00001
40	297	25.6	315	10	AJ346055 shr-00001
41	297	25.6	321	10	AJ345972 shr-00001
42	296	25.6	312	11	F29659
43	295.4	25.5	302	10	AJ346566 shr-00001
44	295.4	25.5	311	10	AJ346339 shr-00001
45	295.4	25.5	314	10	AJ346191 shr-00001

ALIGNMENTS

RESULT 1

LOCUS BF790486 749 bp mRNA 12-JAN-2001

DEFINITION 602250156F1 NIH\_MGC\_81 Homo sapiens cDNA clone IMAGE:4337892 5', mRNA sequence.

ACCESSION BF790486

VERSION BF790486.1 GI:12095536

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 749)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLOM1212 row: o column: 13  
High quality sequence stop: 649.  
Location/Qualifiers 1..749  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4337892"  
/clone\_lib="NIH\_MGC\_81"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: muscle (skeletal); Vector: pDNR-LIB"

FEATURES source



```

QY 1012 gcgtgccccccagccagctaccagcccccctctctctgtgagccggagggtccttaag 1071
Db 155 GCGTGCCCCAGCCAGCCAGCTAGCCAGCCCTCTCTGTGTGTGACGCCGAGGTCCTTAAG 96

QY 1072 aatgctcccgagctaaactagggcccgagctttttctcatgatccaggagcacata 1131
Db 95 AATGGCTCCGGAGCTAACTAGGGCCAGCGCTTTTCTCATGATCCAGGAGCACATA 36

QY 1132 ccacaaactaccacaataaaaaagctg 1158
Db 35 CCACAAACTACCACAATAAAAAAGCTG 9

RESULT 3
BI413513
LOCUS BI413513 875 bp mRNA EST 14-AUG-2001
DEFINITION 602987535F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5143664 5',
RNA sequence.
BI413513
VERSION BI413513.1 GI:15174436
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 875)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11354 row: e column: 09
High quality sequence start: 10
High quality sequence stop: 831.
FEATURES
source
1..875
/organism="Mus musculus"
/strain="C2ECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5143664"
/clone_lib="NCI-CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site:1: NotI; Site:2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'.
TGTACCAACTGTAAGTGGAGCGCCGCTCTGTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 223 a 220 c 272 g 159 t
ORIGIN
Query Match 45.8%; Score 530.2; DB 11; Length 875;
Best Local Similarity 79.7%; Strid. No. 7.8e+109;
Matches 689; Conservative 0; Mismatches 169; Indels 8; Gaps 5;

QY 71 ctggcacagaggaggagaaactccaggagacacacgccagaagctgccatg 130
Db 9 CTTGCCAGGAGAGACTGAGAACTTCGAGATCTGCTCTCTGGGAAGCTGTCCATG 68

```

```

QY 131 gacttgctggtgctgctgagagatgagaagcacacagggggtcagagtgccgctgcagaag 190
Db 69 GACATGCTAGTGTCTAGAGGAAGAGAGAGCCCTCGGGTTCAGAGTCCCTGCTTTACAAAAG 128

QY 191 gtgaaggcccaagagcgtgctgcaagacgtctccctggacctgagcgagagatcatgat 250
Db 129 GTTAGGGCCCAAGAGCGGCTGGCAGACATCCCTGGACCTGGACGTGAGATCATTTGAC 188

QY 251 gtgg---gcgggataccagaacctcatcctcagctgctg---gaagaaacccaagcagaagaagcg 306
Db 189 GTGGTGTGTGGATCCATCAACTCATGAACCTGAGTGAAAAAACCAAGCAGAAAAAGCG 248

QY 307 ggaagctctg-gcggcctcgatgagcgccccccagagcccgagagatcaactggccctg 365
Db 249 AGATGCCCTGTGTGTCAGAGCCAGGAGCCTCTCCAGAGCCAGAGAGATCAACGGCCCTG 308

QY 366 tggatgagagacacctctcctgaaagctcggtgaggggaaaaatgaagctcattgagaagt 425
Db 309 TGAATGAGGAGACATCTCTGAAAGCTGCAGTGGAGGGGAAAATGAAGTCAATTGACAAGT 368

QY 426 tctggtgacggggggtcagccgacacgtgcgaccagtctccgtcgggacagcactgcacc 485
Db 369 ACCTGGCGGAGGAGGTTTCAGCGGACACCTGTGATGAGTTCGCTCGGACAGCACTGCATC 428

QY 486 gagcttccctggaagccacatggaatccttgagaagctctctagataatggggccactg 545
Db 429 GGGCCTCCCTGGAGGACACATGGAGATCTACTGGAGAAAATCTTGAGAGAAATGGGGCCACCG 488

QY 546 tggacttccagatcgctgactgcacagccatcattgggctccgcgcggggccact 605
Db 489 TGGACTTCCAGATCGGCTGGACTGCAGCCATCACTTGGGCCCTGCCGTGGGGGCCACCC 548

QY 606 tagagtggtgaaactcttgcacagccatggagcgacacacacaaatgtgaggggtaagctgc 665
Db 549 TGGAGTGGTGAGACTCTCTGCAAAAGCCGGGGGCCGACACCAACCTGAGAGACAAGCTAC 608

QY 666 tgaacaccccgctgcacg-tggcagtcggacagggcgaggtgagattgtgagagcacttt 724
Db 609 TGAGCACTCCCTGTCATGTTGGCCCTCGCTACTTGACACACGTGGAGATTGTGGAACATTTT 668

QY 725 ctatccctggcctggaatacaatccagagacaggaagggatactgcccctgcataac 784
Db 669 CTCTCCCTGGGCTTGGATATCAATGCCAAAGACACAGAGAGGGGACAGTCCCTCATGAT 728

QY 785 gctgtgaggtcaaccgctacaaatcatcaaaactgctgctcgtcagctggggtgacatg 844
Db 729 GCGGTGAGACTCAACCGCTACAAATCATCAACTGTGCTCTTTGGCATTTGGGGCGGACATG 788

QY 845 atgaccaaagaacctggcagg---aaagaccgcagacgacctggtgcagctcttggcaggctg 902
Db 789 ATGGCTAAGAACTGGCGGGGAAAGAACCCACGGGACCTGGTTCAGCTCTTGCAGCA 848

QY 903 ataccgcgcagccctggagatcc 927
Db 849 GAACCCCGGATGCCCTGGAGCATCC 873

RESULT 4
BE926721
LOCUS BE926721 504 bp mRNA EST 02-OCT-2000
DEFINITION QV2-BT0634-280800-331-b12 BT0634 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE926721
VERSION BE926721.1 GI:10452797
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 504)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare

```

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=6t2-QV2-BT0634-280  
800-331-b12&t3=2000-08-28&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 504.

## FEATURES

source  
1..504  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="BT0634"

/dev\_stage="Adult"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 114 a 147 c 158 g 85 t

## ORIGIN

Query Match 42.5%; Score 492.4; DB 11; Length 504;  
Best Local Similarity 99.8%; Pred. No. 2.2e-100;  
Matches 493; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 555 agatcgctgactgcagcagcattgagctgcctgcgcggggccacttagaggtg 614

DB 11 AGGATCGCTGGACTGCACAGCCATGCAATGGCCCTGCCCGGGGCCACTTAGAGGTGG 70

QY 615 tgaacctctgcaaacccatgagcagacaccaatgtgaggatagaagctgctgagcacc 674

DB 71 TGAACCTTCTGCAAGCCATGGAGCAGACACCAATGTGAGGATAGCTGTGAGCACCC 130

QY 675 cgctgcactggcagtcgccgacagggcaggtgagattgtgagcaactttctacctgg 734

DB 131 CGCTGCACGTGGCAGTCCGGACAGGGCAGGTGGAGATTGTGGATCACTTTCTATCCCTGG 190

QY 735 gcttgaatcaatgccagacagaggaagggatactcctcgatgcgtgagcg 794

DB 191 GCCTGGAATCAATGCCAGACAGGGAAGGGGATGACTGCCCTGATGACGCTGTGAGGC 250

QY 795 tcaaccgctacaaaatcatcaaaactgctcctcgatgggctgacatgatgaccaaga 854

DB 251 TCAACCGCTACAAAATCATCAAACTGCTCCTGATGGGCTGACATGATGACCAAGA 310

QY 855 acctggcagaaagaccccgacggacctgtgacgtctgaggtgataccggcagc 914

DB 311 ACTGTGCAGGAAGACCCCGACCGACCTGGTGCAGCTCTGGCAGGCTGATACCCGSCAG 370

QY 915 ccttgagcatcctgagcggggctgagcataaaggcgtgagggcctaatgatagtg 974

DB 371 CCTTGGAGCATCTTGAGCCGGGGGTGACATTAACGGGCTGAGGGGCCCTAATGATG 430

QY 975 ggcgagagacccctcagcctgtgcccagccagtgatgcgtgcccagccagccagcagcta 1034

DB 431 GCGCAGAGAGCCCTCAGCCCTGTGCCAGCCAGTGAATGCTGCCCCAGCCAGCAGCTA 490

QY 1035 ccagcgcctctctct 1048  
|||||

DB 491 CCAGGCCCTCTCT 504

## RESULT 5

AJ346356 456 bp mRNA EST 14-AUG-2001  
LOCUS SHR-000016-0-I21 HM3/S3 Homo sapiens cDNA, mRNA sequence.  
DEFINITION AJ346356  
ACCESSION AJ346356  
VERSION AJ346356.1 GI:15167539  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 456)  
AUTHORS Laveder,P., De Pitta,C., Toppo,S., Valle,G. and Lanfranchi,G.  
TITLE Specific subtraction of abundant mRNAs in skeletal muscle  
JOURNAL Unpublished (2001)  
COMMENT Contact: Laveder P  
CRIBI Biotechnology Centre  
University of Padua  
Via U. Bassi 58/B, 35121 Padua, Italy  
sequencing centre identifier: HSPD50428  
ABI Chromatograms and other information are available on WWW at  
http://muscle.cribi.unipd.it  
Seq primer: PC2R.

## FEATURES

source  
1..456  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HM3/S3"  
/sex="female"  
/tissue\_type="pectoral muscle (after mastectomy)"  
/note="Vector: pOPD (Custom); Site\_1: EcoRI; Site\_2: NotI;  
caucasian; Skeletal muscle cDNA was depleted of hundred  
most expressed mRNAs through an original two steps  
subtractive procedure"

BASE COUNT 119 a 132 c 126 g 79 t  
ORIGIN

Query Match 37.4%; Score 433.4; DB 10; Length 456;  
Best Local Similarity 99.8%; Pred. No. 3.6e-87;  
Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 724 tctatccctggcctggaaatcaatgccagacaggaaggagatactgcccctgcatga 783

DB 1 TCTATGCCCTGGCCCTGGAAATCAATGCCAGACAGGAAGGGGATACGCCCTGCATGA 60

QY 784 cgctgtgaggtcaaccgctacaaaatcatcaaaactgctcctcgatggggctgacat 843

DB 61 CGCTGTGAGGCTCAACCGCTACAAATCATCAAACTGTGCTCTGATGGGCTGACAT 120

QY 844 gatgaccagaacctggcaggaaagaccccgacgacctgtgagctgtgagcgtga 903

DB 121 GATGACCAAGAACCTGGCAGGAAGAACCCCGACCGACCTGTGTCAGCTGTGAGGCTGA 180

QY 904 taccgagcacgctggagcatcctgagcgggggactaagcataacggcgtgagggcc 963

DB 181 TACCCGGGACGCCCCCTGGAGCATCCTGAGCCGGGGGCTGAGCATACGGGCTGGAGGGGCC 240

QY 964 taatgatgtggcgagagacccctcagcctgtgacgccagtgtaagctgctgcccagc 1023

DB 241 TAATGATAGTGGCGAGAGACCCCTCAGCCTGTGCCAGCCAGTGAATGCGTGCCCCAGC 300

QY 1024 ccagccagctaccagccctctctgtgtgagccggaggggtctctaaagaatggtcccg 1083

DB 301 CCAGCCAGCTACCCAGCCCTCTCTGTGTGTCAGCGGGGGTCTTAAGAAATGGCTCCCG 360

QY 1084 agctaaactgagggccagccctttttctgcatgctccagagacacataccacaactacc 1143

|||||

```

Db 361 ACCTAAGTGGGGCCAGCCCTTTTCTGCGATGATCCAGGAGCAGACATACCACAACTACC 420
Qy 1144 acaataaaagctg 1158
|||||
Db 421 ACAATAAAAGCTG 435

RESULT 6
F24585 435 bp mRNA EST 13-MAY-1999
LOCUS HSPD11011 HM3 Homo sapiens cDNA clone s4000015H03, mRNA sequence.
DEFINITION F24585
ACCESSION F24585
VERSION F24585.1 GI:4810211
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 435)
Lanfranchi, G., Muraro, T., Caldara, F., Pacchioni, B., Pallavicini, A.,
Pandolfo, D., Toppo, S., Trevisan, S., Scarso, S. and Valle, G.
Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
Genome Res. 6 (1), 35-42 (1996)
96276048
Contact: Valle G.
CRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://group.bio.unipd.it.
Location/Qualifiers
1. 435
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="s4000015H03"
/sex="female"
/tissue.type="pectoral muscle (after mastectomy)"
/Note="Vector: pCDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGCGCTCGAGCGCGCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pCDNAII vector."
BASE COUNT 96 a 125 c 141 g 73 t
ORIGIN

Query Match 36.6%; Score 423.8; DB 11; Length 435;
Best Local Similarity 98.4%; Pred. No. 5.1e-85;
Matches 428; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 681 acgtggcgtccggcagggcaggtggagattgtggagcaattctatccctggcctgg 740
|||||
Db 1 ACGTGGCAGTCCGGACAGGGCAGGTGGAGATTGTGGAGAGAGTTCGAGCCCTGGGCGCTGG 60

Qy 741 aaatcaatgccagagcagggagggatactgcctgcctgcctgcctgcctgcctgcctcaacc 800
|||||
Db 61 AAATCAATGCCAGACAGAGGAGGGGATACCTGCTGATGACCTGTGAGGCTCAACC 120

Qy 801 gctacaaatcatcaaaatgctgctcctgcctgcctgcctgcctgcctgcctgcctgcctgcctg 860
|||||
Db 121 GCTACAAATCATCAAACTGCTGCTCCTGCATGGGCTGACATGATGACCAAGAACCTGG 180

Qy 861 caggaaaagaccccgagcactggtgcagctctgcaggtctgatacccgccagccctgg 920
|||||

```

```

Db 181 CAGGAAAGACCCCGAGGACCTGGTGCAGCTCTGGCAGGCTGATACCCGCGACGCCCTGG 240
Qy 921 agcatcctgagcggggtgagcataacggtgagggggtgagggggtgagggggtgagggggtgag 980
|||||
Db 241 AGCATCTGAGCGGGGCTGAGCATAACGGGCTGGAGGGGCTTAATGATAGTGGGGGAG 300
Qy 981 agaccctcagcctgtgcccagccagtgatgctgctgcccagccagccagccagccagccagc 1040
|||||
Db 301 AGACCCCTCAGCCTGTGCCAGCCAGTGAATGCTGCCCGCCAGCCAGCCAGCTACCCAGC 360
Qy 1041 cctctctgtgagcggaggggtcctgaagaatgctgctgagcagcagcagcagcagcagcagc 1100
|||||
Db 361 CCGTCTCTGTGCGAGCGGAGGCTCTAAGATGGCTCCCGGAGCTAACTGAGGGGCCCA 420
Qy 1101 gcttttttctgcat 1115
|||||
Db 421 GCCTTTTCTGCAT 435

RESULT 7
F189893 567 bp mRNA EST 02-NOV-2000
LOCUS 235854 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
DEFINITION F189893
ACCESSION F189893
VERSION F189893.1 GI:11073334
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 567)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, P.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.
and Keele, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 59 row: G column: 21
Seq primer: ATTTAGTGACACATATAG.
FEATURES
Location/Qualifiers
1. 567
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue.type="pooled"
/lab_host="DHL0B"
/Note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 121 a 158 c 195 g 93 t
ORIGIN

Query Match 36.3%; Score 420.4; DB 11; Length 567;
Best Local Similarity 90.7%; Pred. No. 3.1e-84;
Matches 448; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 356 actggcctgtgagtagagagaccttccctgaaagctcggtgaggggaaatgaaggtc 415
|||||
Db 74 ACCGGCCCTGTGGCAGGAGGACATTCCTGAAAGCGCGCTGGAGGGAATGAAGGTC 133

```

QY 416 attgagaatttcctgctgacggggggtcagccagacagctgacacagcttcctgctgagaca 475  
 DB 134 ATTGAGAAGTCTCTGGCGATGGGGTTCCCGCAGACACCTGGACCACTTCGCGCGGACA 193  
 QY 476 gactgcaccagcttcctggaagccacatggaaatccctggagaagcttctagataat 535  
 DB 194 GCCCTGCACGGGCTCCCTGGAGGCGCACATGGAGATCCTGGAGAAGCTTCTGGAGAGT 253  
 QY 536 gggggccactgtgacttcagagatcgctgactgcacagcagcattgggctgcgcgc 595  
 DB 254 GGGGCTACTGTGGACTTCAGATCGGTGGACTCAGACGCGATCATGTGGCGCTGCCGT 313  
 QY 596 gggggccacttaggtgtgaaactctgcaaaagccatgagcagacacacaaatgtgag 655  
 DB 314 GGGGGCCACTTGGAGGTGTGAACTCTGCAAAAGCCGAGGAGCAGACACCACTGTGAGA 373  
 QY 656 gataagctgctgagcaccctgctgacgtggtggtggtggtggtggtggtggtggtggt 715  
 DB 374 GATAAGCTGCTGAGCACCCTGACGTGGCGGTCCGAGCGGGCAGCTGGAGATCGTG 433  
 QY 716 gaggactttctacctgggctgaaatcaatgacagacagagagagagagagagagagag 775  
 DB 434 GAACATTTTCTGTCTGCTGGCGCTGACATCAATGCTCAAAAGCAGAGAGGAGCAGCGCC 493  
 QY 776 ctgcatgacgtgtgaggtcctgacacgctacaaatcaatcaatgctgctgctgctgct 835  
 DB 494 CTGACGACGCGCTGAGACTCAACGCTACAAATCAATCAATCAATCAATCAATCAATCA 553  
 QY 836 gctgacatgatgac 849  
 DB 554 GCTGACATGATGAC 567

RESULT 8  
 F32675  
 LOCUS HSPD25601 417 bp mRNA EST 13-MAY-1999  
 DEFINITION HSPD25601 HM3 Homo sapiens cDNA clone s3000045A02, mRNA sequence.  
 ACCESSION F32675  
 VERSION F32675.1 GI:4818301  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 417)  
 AUTHORS Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,  
 Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.  
 TITLE Identification of 4370 expressed sequence tags from a  
 3'-end-specific cDNA library of human skeletal muscle by DNA  
 sequencing and filter hybridization  
 JOURNAL Genome Res. 6 (1), 35-42 (1996)  
 MEDLINE 96276048  
 COMMENT Contact: Valle G.  
 CRIBI Biotechnology Centre  
 University of Padua  
 Via Trieste 75, 35121 Padua, Italy  
 ABI Chromatograms and other information are available on WWW at  
 http://group.bio.unipd.it  
 Location/Qualifiers  
 1. .417  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="s3000045A02"  
 /clone\_lib="HM3"  
 /sex="female"  
 /tissue\_type="pectoral muscle (after mastectomy)"  
 /note="Vector: pCDNAII (Invitrogen); Site\_1: BstXI;  
 Site\_2: NotI; The library was constructed by G.  
 Lanfranchi. This library is not subtracted nor normalized.  
 The first strand cDNA was primed with a biotinylated  
 oligo-dT-NotI primer  
 (5'-biotin-AACCGGCTCCAGCGCGCGCTTTTCTTTTCTTTT-3'). The  
 ds cDNA was sonicated and size-selected in the range

350-550 bp. The 3' specific fragments were selected by streptavidin coated magnetic beads, ligated to non-palindromic BstXI adapters, NotI digested and directionally cloned into BstXI-NotI cut pCDNAII vector."

BASE COUNT 93 a 124 c 130 g 68 t 2 others  
 ORIGIN

Query Match 35.8%; Score 415; DB 11; Length 417;  
 Best Local Similarity 99.5%; Pred. No. 4.8e-83;  
 Matches 415; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 659 aagctgctgagcaccctgctgacgtggtggtggtggtggtggtggtggtggtggtggtggt 718  
 DB 1 AAGCTGCTGAGCACCCTGCTGACGTGGCAGTCCGCGACAGNGCAGGTGGAGATTGTGGAG 60  
 QY 719 cactttctatccctgggctggtggtggtggtggtggtggtggtggtggtggtggtggtggt 778  
 DB 61 CACTTTCTATCCCTGGGCTGGAAATCAATGCCAGAGACAGGGAAGGGGACTACTCCCTGTG 120  
 QY 779 catgacgtgtgaggtcctgacacgctacaaatcatcaactgctgctgctgctgctgctgct 838  
 DB 121 CATGACGTGTGAGGCTCAACGCTTACAAATCATCAACTGCTGCTCTCTGATGGGCT 180  
 QY 839 gacatgatgac 898  
 DB 181 GACATGATGACCAAGAACCTGGCAGGAAAGACCCCGACGACCTGTGTCAGCTCTGGCAG 240  
 QY 899 gctgatacccgacgctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 958  
 DB 241 GCTGATACCCGCGCCTGAGGATCTGTAGCCGCGGCGCTGAGCATTAACGGCTGGAG 300  
 QY 959 gggcctaagtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1018  
 DB 301 GGGCCTATGATAGTGGCGAGAGACCCCTCAGCCTGTGCCAGCCAGTGAATGCTGCTGCC 360  
 QY 1017 ccagcccgacgctgac 1075  
 DB 361 CCAGCCCGACGCTACCCAGCCCTCTCTGTGTGTCAGCCGCGGAGGCTCTTAAGAATG 417

RESULT 9  
 F36391  
 LOCUS HSPD33981 457 bp mRNA EST 13-MAY-1999  
 DEFINITION HSPD33981 HM3 Homo sapiens cDNA clone sh4-000004-3/E07, mRNA sequence.  
 ACCESSION F36391  
 VERSION F36391.1 GI:4822016  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 457)  
 AUTHORS Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,  
 Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.  
 TITLE Identification of 4370 expressed sequence tags from a  
 3'-end-specific cDNA library of human skeletal muscle by DNA  
 sequencing and filter hybridization  
 JOURNAL Genome Res. 6 (1), 35-42 (1996)  
 MEDLINE 96276048  
 COMMENT Contact: Valle G.  
 CRIBI Biotechnology Centre  
 University of Padua  
 Via Trieste 75, 35121 Padua, Italy  
 ABI Chromatograms and other information are available on WWW at  
 http://group.bio.unipd.it  
 Location/Qualifiers  
 1. .457  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="sh4-000004-3/E07"  
 /clone\_lib="HM3"



/sex="female"  
/tissue\_type="pectoral muscle (after mastectomy)"  
/note="Vector: pcDNAII (Invitrogen); Site.1: BstXI;  
Site.2: NotI; The library was constructed by G.  
Lanfranchi. This library is not subtracted nor normalized.  
The first strand cDNA was primed with a biotinylated  
oligo-dT-NotI primer  
(5'-biotin-AACCGGCTCGAGCGCGCTTTT-3'). The  
ds cDNA was sonicated and size-selected in the range  
350-550 bp. The 3' specific fragments were selected by  
streptavidin coated magnetic beads, ligated to  
non-palindromic BstXI adapters, NotI digested and  
directionally cloned into BstXI-NotI cut pcDNAII vector."  
BASE COUNT 121 a 128 c 126 g 79 t 3 others  
ORIGIN

Query Match 35.4%; Score 410.2; DB 11; Length 457;  
Best Local Similarity 98.8%; Pred. No. 5.8e-82;  
Matches 412; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 741 aaatcaatgcagagacaggaagggtactgacctgcacgtgagcctcaacc 800  
Db 1 AAATCAATGCCAGACAGAGAGGAGGGGATGCTCCCTGCATGACCTGTGAGGCTCAACC 60  
Qy 801 gctcaaaaatcaatgctgctcctgcacgtggggtgacatgatgaccagaacctgg 860  
Db 61 GCTGAAAAATCATCAAACTGCTGCTTCCTGCATGGGCTGACATGATGACCAAGAACCTGG 120  
Qy 861 caggaaagacccccagcagcagcctggtgagcctgacggtgacacccgacccctgg 920  
Db 121 CAGGAAAGACCCCGACGACCTGTGTGACGCTGTGACGAGCTGATACCGGGCACGCGCTGG 180  
Qy 921 agcatcctgagcgggggtgagcagataacggcgtgagggcctaatgatagtgagcag 980  
Db 181 AGCATCTTGAGCGGGGGCTGAGCATACGGCTGGAGGGGCTTAATGATGGCGCAG 240  
Qy 981 agaccctcagcctgtgcccagccagtgatgctgcccagccagcagcagcagcagc 1040  
Db 241 AGACCCCTCAGCCTGTGCGCAGCCCACTGAATGCGTGCCTCCAGCCAGCCAGCTACCCAGC 300  
Qy 1041 cccctctgtgagccgaggggtccctaagaatgctcccgagcagcagcagcagcagc 1100  
Db 301 CCCCTCTGTGTGACGCGGAGGGGCTTAAGAAATGGCTCCCGAGCTTAACGTAGGGGCCCA 360  
Qy 1101 gcctttttctgcatgattccaggagcacataccacaaactaccacataaaaaagct 1157  
Db 361 GCCTTTTCTGCTGATGATCCAGGAGCACATACCACANACTACCANCAATAAAAAAGCT 417

RESULT 10  
AJ346328 408 bp mRNA EST 14-AUG-2001  
LOCUS  
DEFINITION SHR-000016-0-H10 HM3/S3 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AJ346328  
VERSION AJ346328.1 GI:15167511  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 408)  
AUTHORS Laveder, P., De Pitta, C., Toppo, S., Valle, G. and Lanfranchi, G.  
TITLE Specific subtraction of abundant mRNAs in skeletal muscle  
JOURNAL Unpublished (2001)  
COMMENT Contact: Laveder P  
CRIBI Biotechnology Centre  
University of Padua  
Via U. Bassi 58/B, 35121 Padua, Italy  
sequencing centre identifier: HSPD50393  
ABI Chromatograms and other information are available on WWW at  
http://muscle.cribi.unipd.it  
Seq primer: PC2R.

Location/Qualifiers  
1. 408  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HM3/S3"  
/sex="female"  
/tissue\_type="pectoral muscle (after mastectomy)"  
/note="Vector: pOPD (Custom); Site.1: EcoRI; Site.2: NotI;  
caucasian; Skeletal muscle cDNA was depleted of hundred  
most expressed mRNAs through an original two steps  
subtractive procedure"  
BASE COUNT 99 a 125 c 118 g 66 t  
ORIGIN

Query Match 34.7%; Score 402; DB 10; Length 408;  
Best Local Similarity 100.0%; Pred. No. 3.9e-80;  
Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 757 cagggaagggtactgacctgcacgtgagcctcaaccgtcaaaaatcaatcaa 816  
Db 7 CAGGGAAGGGGATGCTCCCTGCATGACCTGTGAGGCTCAACCGCTACAAAATCATCAA 66  
Qy 817 actgtctcctgcacgtggggtgacatgatgaccagaacctggcaggaagaccctgac 876  
Db 67 ACTGTGCTCTCTGCATGGGGCTGACATGATGACCAAGAACCTGGCAGAAAGACCCGAC 126  
Qy 877 ggaacctggtgagcctgagcagcgtgatacccgacgacccctggagcctgagccggg 936  
Db 127 GGACCTGTGTGACGCTGTGCGCAGGCTGATACCGGGCACGCGCTGAGCATCTGAGCGGG 186  
Qy 937 ggtcagcagataacggcgtgaggggcccctaatgatagtgggcgagagacccctcagcctg 996  
Db 187 GCGTACGATACAGGGCTGGAGGGGCTTAATGATAGTGGCGAGAGACCCCTCAGCCTGT 246  
Qy 997 gccagccagtgatgctgctgccccagccagcagcagcagcagcagcagcagcagcagc 1056  
Db 247 GCCAGCCAGTGAATGCGTGCCTCCAGCCAGCCAGCTACCCAGCCCTCTCTGTGTGCGAG 306  
Qy 1057 ccggagggtcctaagaatgctcccgagcagcagcagcagcagcagcagcagcagcagc 1116  
Db 307 CCGGAGGGTCTTAAGAAATGGCTCCCGAGCTTAACGTAGGGGCCAGCGCTTTTCTGCATG 366  
Qy 1117 atccaggagcacataccacaaactaccacataaaaaagctg 1158  
Db 367 ATCCAGGAGCACATACCACAAATACCACATAAAAAAGCTG 408

RESULT 11  
AJ346414 423 bp mRNA EST 14-AUG-2001  
LOCUS  
DEFINITION SHR-000016-0-M09 HM3/S3 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AJ346414  
VERSION AJ346414.1 GI:15167597  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 423)  
AUTHORS Laveder, P., De Pitta, C., Toppo, S., Valle, G. and Lanfranchi, G.  
TITLE Specific subtraction of abundant mRNAs in skeletal muscle  
JOURNAL Unpublished (2001)  
COMMENT Contact: Laveder P  
CRIBI Biotechnology Centre  
University of Padua  
Via U. Bassi 58/B, 35121 Padua, Italy  
sequencing centre identifier: HSPD50512  
ABI Chromatograms and other information are available on WWW at  
http://muscle.cribi.unipd.it  
Seq primer: PC2R.  
Location/Qualifiers  
1. 423

FEATURES  
source

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="HM3/S3"  
 /sex="female"  
 /tissue\_type="pectoral muscle (after mastectomy)"  
 /note="Vector: pOPD (Custom); Site\_1: EcoRI; Site\_2: NotI;  
 caucasian; Skeletal muscle cDNA was depleted of hundred  
 most expressed mRNAs through an original two steps  
 subtractive procedure"  
 BASE COUNT 106 a 125 c 120 g 72 t  
 ORIGIN

Query Match 34.5%; Score 399.2; DB 10; Length 423;  
 Best Local Similarity 99.3%; Pred. No. 1.7e-79;  
 Matches 401; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 755 gacaggaaaggagatacctcctgacgtgacgtgtgagctcaacgcgtacaaaatcattc 814  
 DB 1 GACTGGGAAGGGGATACCTCCCTGCATGACGCTGTGAGGCTCAACCGCTACAAAATCATC 60  
 QY 815 aaactgtctctcgtgagctgacatgatgaccaaagacacgtgcagaaagaccgcg 874  
 DB 61 AAATCTCTCTCTGATGGGCTGACATGCTGACCAAGAACTGCGAGAAAGACCCCG 120  
 QY 875 acggacctgtgagctgtgagctgtatcccgccagccgctgagcatcctctgagccg 934  
 DB 121 ACGGACCTGTGTGAGCTGTGGCAGGCTGATACCGGCGCCCTGTGAGCATCTGAGCCG 180  
 QY 935 ggggctgaacataacgggtgagggcctaataatgtagtgggcagagagacccctcagcct 994  
 DB 181 GGGCTGTGACATAACCGGCTGTGAGGGCCCTAATGATGTGGCGAGAGACCCCTCAGCCT 240  
 QY 995 gtgccagccagtgatgctgctgcccagccagccagctaccagccctctctgtgtgc 1054  
 DB 241 GTGCCAGCCAGTGAATGCTGCCCGAGCCAGCCAGCTACCCAGCCCTCTCTGTGTGC 300  
 QY 1055 agccgaggggtctaagaatgctcccgagctaaactagggcccaagcctttttttgtgca 1114  
 DB 301 AGCCGGAGGGTCTTAAGATGCTCCCGGAGCTAACTGAGGCGCCAGCCCTTTTCTGCA 360  
 QY 1115 tgatccaggagcacataccacaaactaccacataaaaaagctg 1158  
 DB 361 TGATCCAGGAGCATAACCAAACTTACCACAATAAAGAGCTG 404

RESULT 12  
 BE809163  
 LOCUS 537 bp mRNA EST 02-NOV-2000  
 DEFINITION 235852 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
 ACCESSION BE809163  
 VERSION BE809163.1 GI:11073333  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 1 (bases 1 to 537)  
 Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,  
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.  
 and Keeler, J.W.  
 Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 Unpublished (2000)  
 Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel.: 402 762 4366  
 Fax: 402 762 4390  
 Email: smiththe@mail.marc.usda.gov  
 Single pass sequencing. Bases called and alt\_trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.

PCR Primers  
 FORWARD: AGAAACAGCTATGACCAT  
 BACKWARD: GTTTCAGTCACGAGC  
 Plate: 59 row: G column: 19  
 Seq primer: ATTTAGCTGACACTATAG.  
 Location/Qualifiers  
 1. 537

/organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC 2P1G"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
 library made from pooled tissue from testis, ovary,  
 endometrium, hypothalamus, pituitary, and placenta."  
 BASE COUNT 115 a 149 c 186 g 87 t  
 ORIGIN

Query Match 34.0%; Score 393.6; DB 11; Length 537;  
 Best Local Similarity 90.5%; Pred. No. 3.1e-78;  
 Matches 420; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
 QY 356 actggccctgtgtaggagacccctcctgaaagctgcgtgaggggaaatgaaggtc 415  
 DB 74 ACCGGCCCTGTGACGAGGAGACATTCTTGAAGCGCGCTGGAGGGGAAATGAAGGTC 133  
 QY 416 attgagaattcctgctgacgggggtcagcgcacacgtgcacagctccgtcgaca 475  
 DB 134 ATTGAGAAGTTCTGCGGATGGGGTTCCCGACACCTCGACACCTTCCCGCCGACA 193  
 QY 476 gcactgcacccagctccctggaaggccacatcgaaatccttgagagctctctagataat 535  
 DB 194 GCCTTGACCCGGCTTCCCTGGAGGGCCACATGGAGATCCTGGAGAAGCTTCTGGAGAGT 253  
 QY 536 gggggccactgtgagctccaggtgcgtgagctgcacagccatgcattggccctgcgc 595  
 DB 254 GGGGCTACTGTGCTGACCTCCAGGATCGGTGGACTGCACAGCCATGCATTTGGCCTGCCGT 313  
 QY 596 gggggccactgtgaggtgtgaaactctgcaagccatctggaagccatggagcagaccactgtgag 655  
 DB 314 GGGGCCACTTGGAGGTGTGAATCTCTGCAAGCCGAGGAGGACACACAGTGTGAGA 373  
 QY 656 gataagctgtgagcaccgccgtgcacgtggcagctccgacagggcgaggtggagattgtg 715  
 DB 374 GATAAGCTGCTGAGCACCCCACTGACAGTGGCGTCCGACGGGCACTGGAGATCGTG 433  
 QY 716 gagcattctctacccctgggctggaatcaatgcagagagaggggaagatactgcc 775  
 DB 434 GAACATTTTCTGCTCCCTGGGCTGGACATCAATGCCAAAGACAGAGAGGGGACACGCC 493  
 QY 776 ctgcagctgtgaggtgacccgctcaaccgctcaaaaatcatcaaat 819  
 DB 494 CTGCAGGAGCCGTGAGACTCAACCGCTACAAAATCATCAAACT 537

RESULT 13  
 BE809163  
 LOCUS 504 bp mRNA EST 25-APR-2001  
 DEFINITION 214554 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION BE809163  
 VERSION BE809163.1 GI:10240275  
 KEYWORDS EST.  
 SOURCE cow.  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 504)  
 Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,  
 Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett,  
 G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,  
 Perte, G., Holt, I., Karameycheva, S., Liang, F., Quackenbush, J. and

Keele, J.W.  
Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt\_trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCCAGTCAGGACG  
Plate: 71 row: G column: 3  
Seq primer: ATTAGGTGACACTATAG.  
Location/Qualifiers  
1. 504  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone\_lib="MARC 2BOV"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
Library made from pooled tissue from testis, thymus,  
semitendinosus muscle, longissimus muscle, pancreas,  
adrenal, and endometrium."  
BASE COUNT 129 a 127 c 169 g 79 t  
ORIGIN

Query Match 33.7%; Score 390.4; DB 11; Length 504;  
Best Local Similarity 85.9%; Pred. No. 1.6e-77;  
Matches 433; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 15 gcaccatgagagactccgagcggtgcagagggccacagcgtctcatcgagcagcgctgg 74  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 1 GCACAATGGCAGACTCTCAGGAGGTGCAGCGGGCCACGGCACTTATTGAGGAGCGGCTGG 60

Qy 75 cacaggagagaggaatgagaataccgagagagacacgcgcgaagctgccctggact 134  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 61 CACAGGAGGAGGAATGAGAAGTCCCGAGGAACCCACCCTCAGAAGTTGCCCATGGAGA 120

Qy 135 tctgtgtctgagatgagaagcaccaggggtcagagtgacgacctgcagccctgcagaagtga 194  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 121 TcTGTGTGTGGAGGATGAGAAGCACCATTAGGCCTGAGAGTCCGTCCTTACAAAAGGTTA 180

Qy 195 agggccaagagcgctgcgagacgctccctggacctgcgcgaggagatcatcgatgtgg 254  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 181 AGGGCCAGAGCGCGTGGGAAGACATCTTAGACCTCGGAGGAGAGATCATCGAGTGG 240

Qy 255 gggggtccagaacctcatcgagctgcggaagaaacgcgaagcagaagcgagcgctc 314  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 241 GCGGCATCCAGAACCCTCATCCAGCTGCGGAAAAAAGCAAGCAGAAGAAACGGAAGCCC 300

Qy 315 tggcgacctcgatgagcgccccagagcccgagagagatacactggccctgtgatgagg 374  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 301 TGGCGCCTCCAGAGGACCATCTCAGAGCCACAGAGGATTAACGGGCGCTGTGAGCAGAG 360

Qy 375 agacctctgaaagctcggtggagggaataagtgattgagaagttcctggctg 434  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 361 AGACATTTCTGNAAGCTGGGTGGAGGGAGAGATGAAGGTCTATCAGAAAGTTTCTGGCGG 420

Qy 435 acgggggttcagcgacacgctgcgaccagttccgtcgagacacactgcacagagcttccc 494  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 421 ATGGGGGTTCCTCCATCAGACCTGTGATCAGTTTCGCGGAGCTGCCCTGCACCGAGCTCCC 480

Qy 495 tggaggccacatggaatcctgg 518  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
Db 481 TGGAGGCCACATGGAGATCCTCTG 504

## RESULT 14

AJ346604

LOCUS AJ346604 412 bp mRNA EST 14-AUG-2001

DEFINITION SHR-000017-0-115 HM3/S3 Homo sapiens cDNA, mRNA sequence.

ACCESSION AJ346604

VERSION AJ346604.1 GI:15167787

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 412)

Laveder, P., De Pitta, C., Toppo, S., Valle, G. and Lanfranchi, G.

Specific subtraction of abundant mRNAs in skeletal muscle

Unpublished (2001)

Contact: Laveder P

CRIBI Biotechnology Centre

University of Padua

Via U. Bassi 58/B, 35121 Padua, Italy

sequencing centre identifier: HSPD50806

ABI Chromatograms and other information are available on WWW at

http://muscle.cribi.unipd.it

Seq primer: PC2R.

Location/Qualifiers

1. 412

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="HM3/S3"

/sex="female"

/tissue\_type="pectoral muscle (after mastectomy)"

/note="Vector: pOPD (Custom); Site\_1: EcoRI; Site\_2: NotI;

caucasian; Skeletal muscle cDNA was depleted of hundred

most expressed mRNAs through an original two steps

subtractive procedure"

BASE COUNT 98 a 124 c 118 g 71 t 1 others

ORIGIN

Query Match 33.2%; Score 384.4; DB 10; Length 412;

Best Local Similarity 97.8%; Pred. No. 3.5e-76;

Matches 400; Conservative 0; Mismatches 7; Indels 2; Gaps 1;

Qy 750 ccagagacaggaaggagatactgcccctgcagctgagcgtgaggtcgaacccgtacaaa 809

Db 1 CCAGAGNCAGGAGAGGGAT--CTGCCCTGATGACGCTGTGAGGCTCAACCGCTACAAA 58

Qy 810 tcatcaaaactgtctctctgcagtgagggctgacatgatgaccagaacctggcaggaaaga 869

Db 59 TCATCAAACTGTCTCTCTGTCATGGGCTGACATGATGACCAAGAACCTTGGCAGGAAAGA 118

Qy 870 ccccgacgagacctgtgagctctggcagctgatacccgccacgccctggagcatcctg 929

Db 119 CCCCAGCGGACCTGTGTGAGCTCTGGCAGGCTGATACCCGGCACGCCCTGGAGCATCTCG 178

Qy 930 agccgggggctgagcataacggctggaggcctaatgatagtgggcgagagacccctc 989

Db 179 AGCCGGGGCTGAGCATACGGGCTGGAGGGCTTAATGATAGTGGCCGAGACCCCTC 238

Qy 990 agcctgtgccagcccgagtgaaatgcgtgccccagcccgccagcagctaccagccctctg 1049

Db 239 AGCCTGTGCCAGCCAGTGAATGCTGCGCCAGCCAGCCAGCTACCCAGCCCTCTCTG 298

Qy 1050 tctcagccgggggtccttaagaatggctcccgaggactaactgagggccagccttttt 1109

Db 299 TGTGACGCGGAGGGTCTTAAGAATGGCTCCCGGAGCTAACTGAGGGCCAGCCTTTTTT 358

Qy 1110 ctgcatgatccagagacacataccacaaactaccacaaataaaaaactg 1158

Db 359 CTGATGATCCAGGAGCATACCAAACTACCAAACTACCAAACTACCAAACTACCAAA 407

RESULT 15  
BF775625

BF775625	BF775625	489 bp	mrna	EST	25-APR-2001
LOCUS	285833	MARC 3BOV	Bos taurus	cdNA 5', mRNA sequence.	
DEFINITION	BF775625				
ACCESSION	BF775625				
VERSION	BF775625.1	GI:12123525			
KEYWORDS	EST.				
SOURCE	COW.				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.				
REFERENCE	1 (bases 1 to 489)				
AUTHORS	Smith, T.P.L., Grasse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Gasas, E., Wray, J.P., White, J., Cho, J., Fahrenkrug, S.C., Bennett, C.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perteau, G., Holt, I., Karanymcheva, S., Liang, F., Quackenbush, J. and Keele, J.W.				

```

BASE COUNT      112 a      140 c      77 t      0:
ORIGIN
Query Match      33.1%; Score 383.2; DB 11; Length 489;
Best Local Similarity 87.8%; Pred. No. 6.6e-76;
Matches 418; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 557 gatggctggaactgcacagccaatgattggcctgcgcggggggccaattagagtggtg 616
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GACCGGTGGACTGCACAGACCATCACTGGGCCCTGCTGTGGGGGCCACTTGGAGGTGGTG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 617 aaactctgcaaaagccatggagcagacaccaaatgtgaggggataagctgctgagcaccctg 676
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AGACTCTTCGACAGCCGAGGAGCAGACACCAATGTGAGGGATAGACTGCTGAGCACCCCC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 677 ctgcacgtggcagtcctcgacagggcaggtggagattgtggagcaattctatccctgggc 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CTGCATCTGGCAGTCCGACGCGGGCAGGTGGAGATCGTGGAAACATTTTCTATCCCTGGGC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 737 ctgaaatcaatgccagacaggaaggggatactccctgcgatcacgctgtgaagctc 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CTGACATCAATGCCAAGACAGAGAGGGGACAGCGCCCTGCATGATGCTGTGAGGCTC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 797 aaccgctacaaaatcatcaaaactgctctcctgcatgggctgacatgatgaccaagaac 856
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	241	AATCGCTACAAAATCATCAAAATTCCTGCTCTCTGCTATGGGGCTGCATGATGAGCAGAAAC	300
Qy	857	ctggcaggaagaagccccgcagcgacctggtgcagctctctggcaggtctgatacccgccagcc	916
Db	301	CTGGCAGGAAGACCCCCACAGACCTTGGTGCAGCTGTGGCAAGCCGACACCCGGCAGCGT	360
Qy	917	ctggagcatcctgagcgggggctgagcataacggctggaggggcctaataatgattgagg	976
Db	361	CTGAGAACCTTCAGCCAGGGTCAGAGCAGAAATGGACTGGAGGGGTCCACCGAGAGTGGG	420
Qy	977	cgagagacccctcagctctgcagcccgatgaatgcgtgccccagcccgccagc	1032
Db	421	CGGGAGACCCCCAGCGCTCTGGCAGCCGAGTAATGCTGTCCGCGCTTACCAAGC	476

Search completed: April 17, 2002, 09:26:13  
Job time: 4617 sec

---

THIS PAGE BLANK (USPTO)

**THIS PAGE BLANK (USPTO)**